

Db 378 DLDGI 382

RESULT 2

H81831

probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serogroup C)
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: H81831
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, R.; Holtroyd, S.; Javelle, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajadream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: H81831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <P>
A:Cross-references: GB:AL162757; GB:AL157959; NID:97380371; PIDN:CAB85240.1; PID:9738065
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA2021

Query Match 34.5%; Score 814; DB 2; Length 409;
Best Local Similarity 44.4%; Pred. No. 4e-58;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

OY 87 INCSLNDODIMRLACDYLHGEPRAVI-----KTKRSRLDETIMOTI-KGKPGVIOE 140
Db 60 LCCALTDNVTRLACYDKIFPAQLDPSAGQESKAVLNTETVRSLSLDGEAVIVVEK 119
OY 141 TTDPIFLGNENKGMILTKDAKOLEYAAKQFTPLSLSPDLDRNN-TPLWSPRHNPVYLP 199
Db 120 GGDAL-----PADSAGETADITPLSLMDYLDKNDLGLGIVRRHNMVYLMR 166
OY 200 IFMGKPKRSPPTPSH-EAKQFTPNFPAPELKFOVSVKVAEDLWGTDSDFMGYTDQ 258
Db 167 LMVNSPYAPSPGSPRTGTVQEKFGQKRAETKLOVSFSKSKAEDLFTYRADLWGYTOR 226
OY 259 SHMOTFN-GKNSRPRFVHDYOPEIFLTQPVYSDLPMDCKVHIGMAGVHNSGESAKLSR 317
Db 227 SDMOTYNGKRSAPPRNTDYKPEITLTQPVKADLPFGGRLLMLGCFVHQSGQSPRESR 286
OY 318 SMNRAIYLAGMEKMLTYWPMRIWGRIFKEGSGSQDDNDPILDYGYGDPVRELYOLENKS 377
Db 287 SMNRIYAMAGHEMGLTYIPIRVWVRAFDQ-SGDK-NDNPDIADYWGVDVQLQYLRNDRQ 344
OY 378 NISGIVRNPSSGKALQLDVYPLGKISGYFQJFGYGGSLIDYNNHATSFGVGLMLN 437
Db 345 NVYSVLRNPRTGYGAIEAAYTFPIKRLKGVNRGFHGYGESLIDYNNKONGIGTLMFN 404
OY 438 DMWGL 442
Db 405 DLDGI 409

RESULT 3

B36971

outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brook, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm, J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial pldA genes encoding outer membrane phospholipase A3
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:9436880; PIDN:CA54223.1; PID:9436881

A>Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: pldA
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 15.8%; Score 373; DB 2; Length 286;
Best Local Similarity 42.1%; Pred. No. 1.2e-22;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

OY 226 RARPELKFOVSVKVAEDLWGTDSDFMGYTDQSHMQIFNGKNSRFRHNDQPEIFLTQ 285
Db 80 RDEKVKFQSLAFPLMRGILGDNLSLLGASVYQKSWMOUSNSKESAPFRFTNPEOLFGE 139
OY 286 PV-YSDLPMDGKVRMIGMAGVHNSGESAKLSRSMNRAIYLAGMEKMLTYWPMRIWGRIF 344
Db 140 ATDYOPAGH--TLRDIEMQYNHDSNGRSPTRSMNRRLARLMAONGNMLVEVKR---Y 194
OY 345 KEGSGSQDDNDPILDYGYGDPVRELYOLENKSNTSGVRYRPSGKALQLDVYVPLK 404
Db 195 VVGS---TDDNDPIRTKMGYRLKGYQL-CEALISAGQGYMMNTGYGAEIGVSPITK 250
OY 405 GISGYFQJFGYGGSLIDYNNHATSFGVGLMLND 438
Db 251 HVRATQIYSGYESLIDYNNQTRVGVGLMLND 284

RESULT 4

D81279

phospholipase A1 (EC 3.1.1.32) Cj1351 [imported] - Campylobacter jejuni (strain NCTC)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C:Accession: D81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chl, C.W.; Quail, M.; Rajadream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Ba Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: D81279
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <P>
A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CAB73778.1; PID:96966
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: pldA; Cj1351
C:Keywords: carboxylic ester hydrolase

Query Match 15.6%; Score 367; DB 2; Length 329;
Best Local Similarity 33.5%; Pred. No. 4.4e-22;
Matches 106; Conservative 47; Mismatches 125; Indels 38; Gaps 13;

OY 130 IKGRQVYIYQETDPIFLMGNEKMLTKDAKOLEYAAKQFTPLSLSPDLDRNN-TPLM 187
Db 43 LKNSSVLISQEQNNSQATOTONSTITIKKEKQ-----DSRLALANYLGENSESPNPG 96
OY 188 SRPKNPMYVLPITFMHAGPNRSPNTPSHAKQFTPNFPAPELKFOVSVKVAEDLWGT 247
Db 97 IS-SYKMYFLP-FAYSRNSLGVNNKSEA-----KQLDSYKRLFLFNILGL 141
OY 248 DSDLFGYTDQSHMQIFNGKNSRPRVHDYOPEIFLTQPVY-SDLPMDCVRIKMGAVH 306
Db 142 DEKYIAYTQSWMOIY--EHSSPRETNVQGEFIDLPILKDYEFNNLR---VGLH 196
OY 307 HNSGSAAK--LSRSMNRAIYLAGMEKMLTYWPMRIWGRIFKEGSGSQDDNDPILDYGY 364
Db 197 ESNKGSDENLQSRSMNRIYVSFTAILYNNKFLFVPRIMYR---PENKKDDNDPAIILHYGN 253
OY 365 GDVRFYLOLENKSNTSGVRYRNP--SGKALQLDVYVPL-GKISGYFQJFGYGGSLID 421
Db 254 FDNVILAY-LGDDYFTNLMRLKFNHKKGALQVDLGYDIENNGLIYWTLYQIFNGYESLT 312

OY 422 DYNHEATSECVGLMLN 437
||| | : :
Db 313 DYNKHLQRLSTGFLIS 328

RESULT 5
A36971
outer membrane phospholipase A (EC 3.1.1.-) precursor - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: A36971; S40131
J:Brock, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A>Title: Molecular characterization of enterobacterial plid genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: A36971
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-289 <BRO>
A:CROSS-references: GB:X76900; NID:g437024; PIDN:CAA54222.1; PID:g437025
A:Gene: pldA
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 15.2% Score 358; DB 2; Length 289;
Best Local Similarity 39.3%; Pred. No. 2e-21;
Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;

OY 226 RAPELKFQYSVKKAAMDLDGTSDLFMGFYQQSHMOIFNGNSRPFRRVHQDPEFLIQ 285
| : ||||| : : : : | ||||| : : : : | ||||| : : : : |
Db 83 KIDEVFQSLAPLRPLRGILGPNSVLGASTOKSMOLSNKSSESPFRETNEYEPOLFGE 142
OY 286 PV-YSDLPMDGKVRMIGMGAVHHSNGESAKLSRSNNRAYLMAGMEWKNITVPRIWRIGTF 344
143 ANDYRFAGW-TLRHVEMGVNHDSNGSRDSPTSRTMNRLTYRLMAANGNMNLVEWKPPYYI- 199
OY 345 KECSGSQPDNDPDLIDYGGDVAFYLQLENKSNISGTVRRINPRSGKALQLDYVPYLPK 404
Db 200 ----GSTDDNPDIPTKKMYGLOLKIGYHL-GEAYLSAKGOVMNTMGCGAGELSYPMVK 253
||||| | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 405 GISGYFOIFOGYOSLIDYHNHEATSPFGYGLMLND 438
| : : : ||||| | |||||
Db 254 HYRLTYQVYSGIESLIDYFNQTRVGCVGLMLND 287

RESULT 6
PSECAL
phospholipase A1 (EC 3.1.1.32) precursor - Escherichia coli
N:Alternate names: outer-membrane phospholipase A; phosphatidylocholine 1-acylhydrolase;
C:Species: Escherichia coli
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 11-Jun-1999
C:Accession: A22133; S30771; A00771; S66447; F65186
J:Homma, H.; Kobayashi, T.; Chiba, N.; Karasawa, K.; Mizushima, H.; Kudo, I.; Inoue, K.
J. Biochem. 96, 1655-1664, 1984
A>Title: The DNA sequence encoding plid gene, the structural gene for detergent-resistant
A:Reference number: A22133; MUID:85157492
A:Accession: A22133
A:Molecule type: DNA
A:Residues: 1-289 <HON>
A:CROSS-references: GB:X02143; GB:X00780; NID:g42423; PIDN:CAA26081.1; PID:g757840
A>Note: This enzyme is tightly bound to the outer membrane of the cell
R:Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.
Science 257, 771-778, 1992
A>Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 t
A:Reference number: S30660; MUID:92358234
A:Accession: S30711
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-13,'FA',16-289 <DAN>
A:CROSS-references: EMBL:M07049; NID:g836656; PIDN:AAA67617.1; PID:g1488220

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992
R:De Gens, P.; Verheij, H.M.; Riegman, N.H.; Hoekstra, W.P.M.; de Haas, G.H.
EMBO J. 3, 1799-1802, 1984
A>Title: The pro- and mature forms of the E. coli K-12 outer membrane phospholipase A
A:Reference number: A00771; MUID:85003590
A:Accession: A00771
A:Molecule type: DNA
A:Residues: 'MTRQ', 34-289 <DEG>
R:Decker, N.; Merck, K.; Tommassen, J.; Verheij, H.M.
Eur. J. Biochem. 232, 214-219, 1995
A>Title: In vitro folding of *Escherichia coli* outer-membrane phospholipase A.
A:Reference number: S66447; MUID:96048049
A:Accession: S66447
A:Status: preliminary
A:Molecule type: protein
A:Residues: 21-24 <DEK>
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: F65186
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-289 <BIAT>
A:Cross-references: GB:AE000458; GB:U00096; NID:g2367299; PIDN:AAC76824.1; PID:g23673
A:Experimental source: strain K-12, substrain M61655
C:Comment: This enzyme is tightly bound to the outer membrane of the cell.
C:Genetics:
A:Gene: plda
A:Map position: 85 min
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase; membrane bound
F:1-20/Domains: signal sequence #status predicted <SIG>
F:21-289/Product: phospholipase A1 #status predicted <MPT>

```

Query March 15.1%; Score 357; DB 1; Length 289;
Best Local Similarity 39.3%; Pred. No. 2.4e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

OY 226 RAPELKFQVSVYVKAAEDLMGTSDLMFGVYQOOSHMOIFNGKNSRPFVHDYQPEIFLTQ 285
      | : ||| : | : : | : ||| : | : | : | : | : | : | : | : | : | : | : |
DB 83 RKDEVKFLQLSLAFPLMRGILGPNISVLGASVYTKQKSMWQLSNSSESPFETNYEQLFLGF 142
OY 286 PV-YSDLPWDGKVRVIRIGCAVAHHSNGESAKLSRSNNRYILMAGEMKULTYMPRIWGRIF 344
      | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 143 ATDYRFAGM--TLRDYENGYNHDSNGRSDPTSRSNRLYTLRLAMENGMVLVEKPVWYV- 199
OY 345 KEGSGSDPEDNDIIDDYGYGDFRFLYOLEKNSNLSGTVRRNPSPGKALDLYDVPYLGK 404
      ||||| | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 200 -----GNRDDNPDIITKVMGYQLKIGYHL-GDAVLSAAGQINMNTGYGAGELGLSTPYTK 253
OY 405 GISGYFQIFQGYGOSLIDYNHEATSEGVGLMLND 438
      : | : : ||| : ||||| | ||| : |||
DB 254 HVRLTYQVSGGESLIDYNFQTRVGVGMND 287

RESULT 7
886069
outer membrane phospholipase A [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: E86069
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
  11ler, L.; Grobbeck, E.U.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
  Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Accession: AB85480; MUID:21074935; PMID:11206551
A:Reference: E86069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <STO>

```

A:Cross-references: GB:AE005174; NID:q12518695; PIDN:AG59017.1; GSPDB:GN0145; UWGP:Z53
A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:
A:Gene: plda

C:Superfamily: bacterial phospholipase A1

Query Match 15.1%; Score 357; DB 2; Length 289;
Best Local Similarity 39.3%; Pred. No. 2.4e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

OY 226 RAPELKFQVSVYVKAEDLMDGTDSDLMFGYTOQSHMOIFENGNSRPFRVHDQPEIFLTQ 285
DB 83 RKDEKRFQSLAFPLMRGILGPNLSVGLASYTKSMWOLSNSESSPFRETNTIEPOLIGF 142
OY 286 PV-YSDLPMDGKVRMIGGAVHNSGESAKLSRSNRRAYLMAEMKMLTVMPRTIGRTF 344
DB 143 ATDYRFAGM--TLRDVEMGYNHDSNGRSDPTSRSMNRLYTRLMAENGWMLVEKKPMYYV- 199
OY 345 KEGSSQDDNDLIDYGYDVRFLYOLENKSNSIGVRRNPRSGKALDLDYVPLGK 404
DB 200 -----GNTDDNDPDIRKMGYVOLKIGYHL-GDAVLAKGYNMNTGYGAEGLGSLPYTK 253
OY 405 GTSGYFOIFQYGOSLIDYVNEHATSFYGLMLND 438
DB 254 HVRLTYVSGSGESLIDYVNEHATSFYGLMLND 287

RESULT 8
C36971
outer membrane phospholipase A (EC 3.1.1.-) precursor - Proteus vulgaris

C:Species: Proteus vulgaris
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999

C:Accession: C36971; S40130
R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm

J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial plda genes encoding outer membran

A:Reference number: A36971; MUID:94131966
A:Accession: C36971

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-289

A:Cross-references: EMBL:X76902; NID:q436889; PIDN:CAA54224.1; PID:q436890

A:Note: authors translated the codon GAG for residue 74 as Gly and CGA for residue 115 a

C:Genetics:
A:Gene: plda

C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 13.6%; Score 321.5; DB 2; Length 289;
Best Local Similarity 33.3%; Pred. No. 1.8e-18;
Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

OY 179 LDRNTPRLMSSRPHPMVLPIFMHGKPNRSPNTPSHEAKQFTPEFRAPELKFPQSVKV 238
DB 43 LOEHNP-FTLYPVSNLLTY-----TSDLNKKAISYMSDNA-NKDEKFLQSLAF 95

OY 239 KAEDLMGTDSDLMFGYTOQSHMOIFENGNSRPFRVHDQPEIFLTQPV-YSDLPMDK 297
DB 96 PLMRILDLNLSGLASYTKSMWOLSNSESSPFRETNTIEPOLIGFATDVSQGM--TL 153

OY 298 RMIGGAVHNSGESAKLSRSNRRAYLMAEMKMLTVMPRTIGRTFEGSGSQDDNDP 357
DB 154 RDAEFGYVHNSGSDPTSRSMNRLYTRLMAENGWMLVEKKPMYYV-----GDTSSKN 207

OY 358 ILDYGVDVRFYOLENKSNSIGVRRNPRSGKALDLDYVPLGKISGYFOFGY 417
DB 208 ITKWTGYQLIGYOL-GEAVLSAKGYNMNTGYGAEGLGSLPYTKHVRFTQVSGYG 266

OY 418 OSLLIDYVNEHATSFYGLMLND 438
DB 267 ESLIDYVNEHATSFYGLMLND 287

RESULT 9
C64582
phospholipase A1 precursor - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: C64582
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Meldman, J.M.; Fujii, C.; Bowman, C.; Matthey,

Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpek, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467

A:Accession: C64582
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-355 <TOM>
A:Cross-references: GB:AE000564; GB:AE000511; NID:q2313602; PIDN:AAD07564.1; PID:q231

Query Match 10.4%; Score 246.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 2.8e-12;
Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

OY 157 KDKAQLEYAAKQPTLSLFDLDRNTPRLMSSRPHPMVLPIFMHGKPNRSPNPSHE 216
DB 69 KKYLMMDYLGTFYFLPFYHSF-----TPIFQWYHPNINP----- 102

OY 217 AKQTPNFRAPELKFOVSVYVKAEDLMDGTDSDLMFGYTOQSHMOIFENGNSRPFRVHD 276
DB 103 ---YORNEF-----KFQLSFRVPRFRHILWTKGTYLAVTQTDWFOYINDPOSAPPRMKN 154

OY 277 YQPELFTQPYVSDLPMDKGV---RMIGGAVHNSG--ESAKLSRSNRRAYLMAEMKKN 332
DB 155 FMPELIYVYPI-NKRPFGKIGNSEIWMIGMHSNGVAGACQYPPFK-----EENPENQ 209

OY 333 LTVMPRI-----WGRIFEGSGSQP-----DDNDP 357
DB 210 FPGQPIVIVKDYNGKQDVWGGCRSVSAGQRPFLRWMEKGLKINVAWYVPYVDOSNPN 269

OY 358 ILDYGVDVRFYOLENKSNSIGVRRNPRSGKALDLDYVPLGKISGYF 410
DB 270 LIDYGYNAKIDYRRGRHHEFLQYDIFTQYWRD--RMHGAFLGYTRINPFVGIYA 327

OY 411 OIFQYGOSLIDYVNEHATSFYGLMLND 437
DB 328 QMFNGYGDLYEYDVFSNRIGYIRLN 354

RESULT 10
H71930
probable phospholipase a1 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori
A:Variety: Strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: H71930

R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.

Yves, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F

Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557
A:Accession: H71930

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-355 <ARN>
A:Cross-references: GB:AE001479; GB:AE001439; NID:q4154979; PIDN:AAD06029.1; PID:q415

A:Experimental source: strain J99
C:Genetics:
A:Gene: plda

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 8.6e-12;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

157 KKDAKQLEYAAKQPTPLSLFSDLDNRNMTPLWSSRPHNPMYVLPITMGKPKNSPPTPSHE 216
 Db KKLMMMDYLGTFLPEVHSF-----PPIQWYPRNP----- 102

217 AKQFPNFRAPBELFQVSVKKAADLMGTDSDLMFGTQOSHMOIFNGKSRPRVHD 276
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

277 YQPELFLTOPVYSDLPMDQKV---RMIGKGAVHSNG--ESAKLSSRNNR----- 321
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

155 FMPBELIYYPI-NKFRPGKIGNFSEIWMQHISNGVAGACQOPFKNEGPNEMQFQ 213
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

322 -----AYLMAG-----MEMK--NLTVMRPIGRIFKEGSGSQPD- 354
 Db PVIVKDYNGQKDVWGGCRSVSAGNALCFVLWMEKGLKIMVAMPYV-----PYDQ 265

214 PVIVKDYNGQKDVWGGCRSVSAGNALCFVLWMEKGLKIMVAMPYV-----PYDQ 265
 Db PVIVKDYNGQKDVWGGCRSVSAGNALCFVLWMEKGLKIMVAMPYV-----PYDQ 265

355 -NPOLDIYGYGVDFVFLY-----QLENKSNISGVKRNPRSGKALOLDIYVPLGKI 406
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

266 SNPOLDIYGYGVDFVFLY-----QLENKSNISGVKRNPRSGKALOLDIYVPLGKI 406
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

407 SGYFOIFQGYGOSLIDYNEHATSPFGVGLMLN 437
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

324 GYVQWFMNGYGDGLXEDYVFSNRIGVGRIN 354
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

RESULT 11
 TYVTRR
 protein kinase (EC 2.7.1.37) raf - rat
 N.Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-8
 C.Species: Rattus norvegicus (Norway rat)
 C.Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
 C.Accession: B26126
 R.Ishikawa, F.; Nagao, M.; Sugimura, T.
 Mol. Cell. Biol. 7, 1226-1232, 1987
 A.Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein.
 A.Reference number: A26126; MUID:87172791
 A.Accession: B26126
 A.Molecule type: mRNA
 A.Residues: 1-602 <ISH>
 A.Cross-references: GB:M55428; NID:g206546; PIDN:AAA42002.1; PID:g206547
 C.Genetics:
 A.Gene: raf
 C:Superfamily: rat protein kinase raf; protein kinase homology
 C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonine
 F:301-567/Domain: protein kinase homology <KIN>
 F:309-317/Region: protein kinase ATP-binding motif
 F:329/Active site: Lys #status predicted
 F:453/Binding site: phosphate (ser) (covalent) (by autophosphorylation) #status predicted

Query Match 4.8%; Score 112.5; DB 1; Length 602;
 Best Local Similarity 19.3%; Pred. No. 0.43;
 Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;

23 QAOAVPNPVAFYDEVRENDLQDNELPIDVOSATQASDTANPLDEHEPELYTTALEN 82
 Db ELKLNNOVKVLETKRKELETAQDRMLGI-----OSQFTRAKKELEAKRDLIRT---N 131

81 ELKLNNOVKVLETKRKELETAQDRMLGI-----OSQFTRAKKELEAKRDLIRT---N 131
 Db ELKLNNOVKVLETKRKELETAQDRMLGI-----OSQFTRAKKELEAKRDLIRT---N 131

83 KMLNLCSALNDIMRLACVDTLVHGEFPAVITKRSITLDETIQ--TIKGRPOVIYOE 140
 Db ERLSOEVEYLTEDVKRL--NEKLKESNT---TKGELQLLDELQASDVVYKREKLEOE 186

141 -----TTPDIFLMGNKNG-----MLTKDAKQLEYAAKQPTPLS 174
 Db KELLNNOVKVLETKRKELETAQDRMLGI-----OSQFTRAKKELEAKRDLIRT---N 131

187 KELLNNOVKVLETKRKELETAQDRMLGI-----OSQFTRAKKELEAKRDLIRT---N 131
 Db KELLNNOVKVLETKRKELETAQDRMLGI-----OSQFTRAKKELEAKRDLIRT---N 131

175 LSFDDLDRNNTPLMSRPHNPMYVLPITMGKPKNSPPTPSHEAKQFTPNFRAPBELIKOV 234
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

247 LS-SSPNNLSPTGWSQPKTP-----VPAQREBARAFGSGTOEKNKIRPRGQDSVYWEI 298
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

235 SVKVAEDLMGTDSDLMFGTQOSHMOIFNGKSRPRVHDQPEI----- 281
 Db EASEVALSTRISGS--FGTYVKKGMHGDVAVKILKVDPPKQVLOAEIRNEVAVLRKTR 355

282 -----FLTQ-----PYVSDU--PMDGKVRMI-----GMGAVHR 307
 Db HYNILLFPGYTKRDNALITVOMCEGSSLYKHLHYOETKQOMQLDIANQTOAGMDYLLHA 415

308 SNGESAKLSRNNRAYLMAGMEWK---NLVMPRIWGRIFKEGSGSQ----- 351
 Db KNIHRDMKS--NNIFLHGLTVKIDFGLATVKSWM-----SGSQVEQPTGSLVM 466

416 KNIHRDMKS--NNIFLHGLTVKIDFGLATVKSWM-----SGSQVEQPTGSLVM 466
 Db KNIHRDMKS--NNIFLHGLTVKIDFGLATVKSWM-----SGSQVEQPTGSLVM 466

352 -----PDNDPDL--DYVGYGVDFVFLY-----QLENKSNISGVKRNPRSGKALOLDIYVPL 402
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

467 APEVIRMODNNPSPFOSDVSSTGYV--LYEL-----MTGELPYSHINNDDI---IFMV 515
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

403 GRG 405
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

516 GRG 518
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

RESULT 12
 A29003
 cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
 N.Alternate names: endo-1,4-beta-glucanase
 C.Species: Bacillus sp.
 C.Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2000
 C.Accession: A29003
 R.Fukumori, F.; Kudo, T.; Narahashi, Y.; Horikoshi, K.
 J. Gen. Microbiol. 132, 2329-2335, 1986
 A.Title: Molecular cloning and nucleotide sequence of the alkaline cellulase gene fro
 A.Reference number: A29003; MUID:87085443
 A.Accession: A29003
 A.Molecule type: DNA
 A.Residues: 1-800 <FUK>
 A.Cross-references: GB:D00066; GB:N00066; NID:g216223; PIDN:BAA00045.1; PID:g216224
 A.Experimental source: strain 1139
 A.Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
 A:Pathway: cellulose degradation
 C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
 C:Keywords: glycosidase; hydrolysis; polysaccharide degradation
 F:585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.7%; Score 110.5; DB 2; Length 800;
 Best Local Similarity 20.0%; Pred. No. 0.96;
 Matches 96; Conservative 53; Mismatches 163; Indels 169; Gaps 25;

9 TILSLSCFALIAQOAVPNPVAFYDEVREND---LQDNELPIDVOSATQASDTA 65
 Db TQOLISSILILV-LTSLPPTALAEGRNREDFKHLGNDVKKRPSGALQLOEVDQ 65

66 NPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMLACY-DLVHGETP 111
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

66 NPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMLACY-DLVHGETP 111
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

126 ELIKSRVIGIDILAEINDMYVIVDMVHAQDPDPAVGAEDFFRIDALYPPNNHITY 185
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

139 QETDP-----IFLMGKGMVLTKKDAKQLEYAAKQPTPLSDDDRN---NTPLM 187
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

186 ELANEPSSNNNGAGIPNNEEGNAV---EYADPIYEMLRDSCNADNIIIVGSPWM 240
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

188 SSRP-----HNMPVLPDIF--MHGKPKNS--PNTPSHEAKQFTPNFRAPBELIKOV 233
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

241 SQRPDLAANPIDDHTMTTIVHYGSHAASTESYPERFPNBERGVMNTRYA-----LE 296
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

234 VSVKVAEDLMGTDSDLMFGTQOSHMOIFNGKSRPRVHDQPEI----- 281
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

297 NGVAVATE--WGTSGANGDGPYFDEADVWIEFLNENNISWANWISLTN-KNEVSGAFYP 353
 Db ---YORNEP-----KFOISFRVPRFRHILMTGTLYLATOTNMFOIYVDPQAPARMIN 154

R:McMurray, A
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19225
A:Accession: T20122
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <WT>
A:Cross-references: EMBL:Z270750.; PIDN:CAA94750.1.; GSPDB:GN00023.; CESP:F25B3.1
A:Experimental source: clone C50F4
R:Gardner, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19406
A:Accession: T21324
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <WT>
A:Cross-references: EMBL:Z270752.; PIDN:CAA94758.1.; GSPDB:GN00023.; CESP:F25B3.1
A:Experimental source: clone F25B3
C:Genetics:
A:Gene: CESP:F25B3.1
A:Map position: 5
A:Introns: 25/3.; 47/2.; 199/1.; 242/1.; 315/2.; 544/3.; 600/1.; 626/1.; 646/1.; 666/3.; 706/3.; 75

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:34 ; Search time 44.45 seconds

(without alignments)
757,461 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSLSCFAILA.....YNHEATSRGVLMLNDMGL 442

Scoring table:

BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_68:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	EB1195
2	815	34.5	409	2	H81831
3	373	15.8	286	2	B36971
4	366	15.5	329	2	D81279
5	358	15.2	289	2	A36971
6	357	15.1	289	1	PSECA1
7	357	15.1	289	1	E86069
8	321.5	13.6	289	2	C36971
9	246.5	10.4	355	2	C64582
10	240.5	10.2	355	2	H71930
11	110.5	4.7	602	1	TVPTRR
12	109.5	4.6	800	2	A29003
13	109.5	4.6	822	2	JT0611
14	108	4.6	824	3	JC7532
15	105.5	4.5	901	2	T20122
16	103	4.4	783	2	JC5467
17	101	4.3	719	2	A42893
18	101	4.3	4307	2	T20721
19	99.5	4.2	1658	2	T42642
20	99	4.2	719	2	S28031
21	99	4.2	719	2	S28033
22	99	4.2	5005	2	F82884
23	98.5	4.2	660	2	E83656
24	98.5	4.2	810	2	S49744
25	98	4.2	719	2	S28034
26	98	4.2	719	2	S28032
27	97.5	4.1	1478	2	S20117
28	96.5	4.1	797	2	T46737
29	96.5	4.1	838	2	A96557

30	96	4.1	608	2	S28036	penicillin-binding
31	95	4.0	719	2	S28035	penicillin-binding
32	95	4.0	719	2	S28038	penicillin-binding
33	95	4.0	719	2	S28037	penicillin-binding
34	95	4.0	765	2	T35719	chitinase - Strept
35	94.5	4.0	397	2	A35136	cellulase (EC 3.2.
36	94.5	4.0	601	2	T26062	hypothetical prote
37	94.5	4.0	655	2	T26061	hypothetical prote
38	94	4.0	1115	2	D69517	DNA polymerase III
39	93.5	4.0	564	2	T40777	feric reductase t
40	93	3.9	791	2	H96839	hypothetical prote
41	92.5	3.9	857	1	A41369	S-receptor kinase
42	92.5	3.9	1379	1	S01254	hepatocyte growth
43	92	3.9	324	2	B69521	hypothetical prote
44	92	3.9	888	2	S50801	AMP deaminase homo
45	91.5	3.9	486	2	S30959	gene 14 protein -

ALIGNMENTS

Query Match	Best Local Similarity	Score	DB 2:	Length	382:
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;	
87	INCSALNODIMRLACYPDLVHGETPAVY-----KTKRSIRLDETIMQTI-KGKPOVYOE	140			
33	LOCALALDNTVTRLACVIRFAOQLPSSAGQEGSKAVLNITEVRSLSKGEAVIYER	92			
141	TTDFIFLMNGEKMGLTKKAKOLEYAAKQFTPLSLSFDDLRRN-TPLMSSRPHNPVLP	199			
93	GGDAL-----PADSAGETADYITPLSLMDLKDNDLGLGVEHNPVIMP	139			
200	IFMNGKRNRSPTNPFSHAR-QFTNEFRAPELKQVSVYKKAADMDLGTSDMLFGTQQ	258			
140	LMVNNNSPNVAPSGPTKCTTQVEKFGQKRAETKIQVSEKSLADLDKTRADLFEVGTOR	199			
259	SHMOIFR-GKNSRPFRRHDOPEIFLTPYVSDLPWGQKRMIMGAVHNSNGESATLSR	317			
200	SDMOIYOGKRSAPFRNTDYKPELFLQPYKADLPFGGRMLMGAGVHOSNGOSRPSR	259			
318	SWNRAYLMAEMKMLTVMPRIMGRIFKEGSGQPDNDPILDYVYGADVRFYOLENKS	377			
260	SWNRIVYAMAGEMKGLVIVRVWVAFDQ-SGDR-NDNPIADVMGVDKLYQRLNDRQ	317			
378	NISGTAVYRNRSKGALQDLYVPLGKIGISGYROIPOGCGOSLIDYVHEATSRGVGLMLN	437			
318	NYSVLEKYNPKTGALAAVYTPPIKGLKGVYVGFHGESESLIDYVHNKONGIGILMFN	377			
438	DMWGL 442				

Db 378 DLDGI 382

RESULT 2

H81831

probable phospholipase A2 [imported] - Neisseria meningitidis (strain 22491 serogroup H81831)
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: H81831
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajadream, Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.
 A:Reference number: A81775; MUID:20222556
 A:Accession: H81831
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-409 <PAR>
 A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738065
 A:Experimental source: serogroup A, strain 22491
 C:Genetics:
 A:Gene: NMA2021

Query Match 34.5%; Score 815; DB 2; Length 409;
 Best Local Similarity 44.7%; Pred. No. 2,1e-58;
 Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

87 INCSLNDINDIACDPLVHGEFPAVI-----KTKRSIRLDETIMORT-KGKPOVYVOE 140
 Db 60 LCCALDNDVTRTACDTRIFFAQQLPSSAGQESKAVNLTEYTRSSLDKGEAVIVVEK 119
 QY 141 TDPFELMGNEKMLTKDAKQLEAQAQFPLSLFDLDRNN-TPLWSSRHPNMYVLP 199
 Db 120 GGDAL-----PADSAGETADYPLSLMYDLDKNDLGLLGVRRHNNMYLMP 166
 QY 200 IFMHKRPNSPPTPSHEAR-QTPNEFRAPLEFOYSKVKVAEDLWGTDSPLMGTYTOO 258
 Db 167 LMVNSPVPAPSPPTGTTVOEKFGQKRAETKLQVSFSKJAEDLFFKRAIDLMEGYTOR 226
 QY 259 SHMOJFN-GKNRSPRVHDYOEIFLTQPVYSDLPMDGKVRMIGAVHHSSESAKLSR 317
 Db 227 SMOJYINGKRSAPFRNDYKREIFLTQPVKADLPFGGRLMLGAFVHQSGSRPESR 286
 QY 318 SMNRAVYLAGMEKMLTYVPRIMGRIFKESGSDPDNDPILDYGYGDPVRLYOLENKS 377
 Db 287 SMNRIYAGMEKMLTYVPRIMGRIFKESGSDPDNDPILDYGYGDPVRLYOLENKS 344
 QY 378 NISGIVRINPRSGKALQDLYVPLKGISGTFQJFQJGOSLIDYNHGATSFYGLMLN 437
 Db 345 NYVSLRINPKTGALIEAAYTFPIKGLKGVRRGFHGESLIDYNHKONGIGLMEFN 404
 QY 438 DWML 442
 Db 405 DLDGI 409

RESULT 3

B36971

outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
 C:Species: Klebsiella pneumoniae
 C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
 C:Accession: B36971; S40129
 R:Brick, R.G.; P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm, J. Bacteriol. 176, 861-870, 1994
 A:Title: Molecular characterization of enterobacterial plid genes encoding outer membrane phospholipase A2
 A:Reference number: A36971; MUID:94131966
 A:Accession: B36971
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <BRO>
 A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881

A:Note: authors translated the codon AAG for residue 112 as Arg
 C:Genetics:
 A:Gene: plidA
 C:Superfamily: bacterial phospholipase A1
 C:Keywords: carboxylic ester hydrolase

Query Match 15.8%; Score 373; DB 2; Length 286;
 Best Local Similarity 42.1%; Pred. No. 9.6e-23;
 Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

QY 226 RARLEKFOYSKVKVAEDLMGTDSOLMGTYTOOSHWQFNGKNSRPRVNDQPEIFLTQ 285
 Db 80 RDKVKFQSLIAFPMLRGILGDSNLGASYYTKSWWQSLNSKESAPFRETYNEPOLFLGF 139
 QY 286 PV-YSDLPMDGKVRMIGAVHHSNGESAKLSRSNRAVYLAGMEKMLTYVPRIMGRIF 344
 Db 140 ATDYOPACH--TLRDIEMKYNHDSNGRSDPYSRNNRLYARLMAONGMNLVYKRW---Y 194
 QY 345 KESGSDPDNDPILDYGYGDPVRLYOLENKSNTSGVRYNPRSGKALQDLYVPLK 404
 Db 195 VVGS--TDDNDPILKMGYRLKVGYL-GEALISAGQVMMNTGYGAEIVGSPITK 250
 QY 405 GISGTFQJGOSLIDYNHGATSFYGLMLND 438
 Db 251 HVRATQIYSGGESLIDYNQTRVGVGLMND 284

RESULT 4

D81279

phospholipase A1 (EC 3.1.1.32) Cj1351 [imported] - Campylobacter jejuni (strain NCTC D81279)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
 C:Accession: D81279
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chli, C.W.; Quail, M.; Rajadream, M.A.; Rutherford, K.M.; Vanyilet, A.; Whitehead, S.; Ba, Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals A:Reference number: A81250; MUID:20150912
 A:Accession: D81279
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-329 <PAR>
 A:Cross-references: GB:AL130708; GB:AL111168; NID:g6968723; PIDN:CAB37778.1; PID:g6966
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: plidA; Cj1351
 C:Keywords: carboxylic ester hydrolase

Query Match 15.5%; Score 366; DB 2; Length 329;
 Best Local Similarity 33.2%; Pred. No. 4.4e-22;
 Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

QY 130 IKGRPOVYVOETDPIFLMGNEKMLTKDAKQLEAQAQFPLSLFDLDRNN--PLM 187
 Db 43 LKNSVLLISQEQNNSQATQTONSTIKKEKQ-----DSRLALANYLGENSEFNPVG 96
 QY 188 SSRPNPMVPLPIFMHGRPNRSPPTPSHEARQFTNEFRAPLEKFOYSKVKVAEDLMGT 247
 Db 97 IS-SYKMYFLP-FAYSRNSLGVNNKSEA-----KQLQSVKRLFLFNILGL 141
 QY 248 DSDLMFGTYTOOSHWQIFNGKNSRPRVHDYOEIFLTQPVY-SDLPMGKVRMIGAVH 306
 Db 142 DEKYIATQNSWMOIY--EHSSPRETYNQEPIIDLPLKYKDEFFNNLR---VGLH 196
 QY 307 HSNESAK--LSRSNRAVYLAGMEKMLTYVPRIMGRIFKESGSDPDNDPILDYGY 364
 Db 197 ESNKGDENLQSRNRYIVSTAILYNNKFLVPRIMGRIFKESGSDPDNDPILDYGY 253
 QY 365 GVPVRLYOLENKSNTSGVRYNPR--SGKALQDLYVPL-KGISGTFQJGOSLIDY 421
 Db 254 PDVNLAY-LGDDYFINTMLRNRLKFNHKKGALQVDLGDIFNNGIYWLQYFNGYGESLI 312

OY 295 GAVRMIGCAVHHSNGES-----AKLSRSMRAYLMAEMKNTLWPRIMG 341
 Db 411 GTRKQGFV-----NDSPEDVIEINAGALKSLGDSNDVSEGNWYANARLSADGWG 464
 OY 342 R 342
 Db 465 K 465

RESULT 13 J70611

cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
 N:Alternate names: endo-1,4-beta-glucanase
 C:Species: Bacillus sp.
 C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
 C:Accession: J70611
 R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
 Biosci. Biotechnol. Biochem. 56, 872-877, 1992
 A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
 A:Reference number: J70611; MUID:92305459
 A:Accession: J70611
 A:Molecule type: DNA
 A:Residues: 1-822 <SUM>
 A:Cross-references: GB:M84963; NID:9289264; PID:NAA73189.1; PID:9289266
 C:Function:
 A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
 A:Pathway: cellulose degradation
 C:Superfamily: Bacillus sp. alkaline cellulase; Thermocoga xylanase A amino-terminal ref
 C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
 F:585726/Domain: Thermocoga xylanase A amino-terminal repeat homology <TXM>

Query Match 4.6%; Score 109.5; DB 2; Length 822;
 Best Local Similarity 19.8%; Pred. No. 1.2;
 Matches 95; Conservative 54; Mismatches 163; Indels 169; Gaps 25;

OY 9 TLTSLSCFAIIAQAKVNPVAFVDEVRSEND---LGQNDHELPIDVQSAQSASTORA 65
 Db 7 TKQLISSIILVL-LTSLFPALAEAGNTREDNFKHLGNDVKKRSEAGALQLOEVDQ 65
 OY 66 NPL-DEHEPELY-----TTALENKTMLINCAL-----NODIMRLACY-DTLVHGEP 111
 Db 66 MTLVDQHGKIDLRGMSTHGLQWPEIILNDNAYKALANDMESNMIRLAMYGENGYASNP 125
 OY 112 AVIKTKRSIRIDETI-----WQT-----IKGRPOVY 138
 Db 126 ELIKSRVIGIDLALENDMYIVDWHVHAPGDPDPVYAGADEFRDIALALYPPNPHIT 185
 OY 139 QETTPD-----IFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDDLRN---NTPLW 187
 Db 186 ELANEPSSNNNGAGIPNNEEGMAVK-----ETADPIVEMLRDGNADNDNIIIVGSPW 240
 OY 188 SSRP-----HNPMYVLPF--MHGKPNRS--PNTPSHEARQFTPNERRAPELKRQ 233
 Db 241 SGRPLADNPIDIDHHTMYTVHFTGSHAASTESTYPPETPNSEKGNVMSNTYA---LE 296
 OY 234 VSVKKAADLMGT-----DSDLMFGYTOO-----SHWQIFNGKNS-----RP 271
 Db 297 NCVAVFATE--WGTQASGDGPGYFDEADVWIEPLNENNISMANNSLTN-KNEVSGAFTP 353
 OY 272 FRVH-----DYQEIFLTQPVYSDLPW---D 294
 Db 354 FELGSMNTSLDPGPDQVWPEELSLSGEYVRAIKGVNTEP---IDRKTKYTKVLMDFND 410
 OY 295 GAVRMIGCAVHHSNGES-----AKLSRSMRAYLMAEMKNTLWPRIMG 341
 Db 411 GTRKQGFV-----NDSPEDVIEINAGALKSLGDSNDVSEGNWYANARLSADGWG 464
 OY 342 R 342
 Db 465 K 465

RESULT 14
 JC7532

cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-S237)
 N:Alternate names: endo-1,4-beta-glucanase
 C:Species: Bacillus sp. (strain KSM-S237)
 C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
 C:Accession: JC7532; PC7107
 R:Hakamada, Y.; Hatada, Y.; Kolke, K.; Yoshimatsu, T.; Kawai, S.; Kobayashi, T.; Ito,
 Biosci. Biotechnol. Biochem. 64, 2281-2289, 2000
 A:Title: Deduced amino acid sequence and possible catalytic residues of a thermostabl
 A:Reference number: JC7532; MUID:21036886
 A:Accession: JC7532
 A:Molecule type: DNA
 A:Residues: 1-824 <HAK>
 A:Cross-references: DDBJ:AB018420
 A:Experimental source: strain KSM-S237
 A:Accession: PC7107
 A:Molecule type: protein
 A:Residues: 31-50 <HAK>
 C:Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent
 C:Genetics:
 A:Gene: Egl-237
 C:Keywords: hydrolase; glycosidase

Query Match 4.6%; Score 108; DB 3; Length 824;
 Best Local Similarity 19.8%; Pred. No. 1.5;
 Matches 95; Conservative 60; Mismatches 159; Indels 166; Gaps 28;

OY 9 TLTSLSCFAIIAQAKVNPVAFVDEVRSEND---LGQNDHELPIDVQSAQSASTORA 65
 Db 7 TKQLISSIILVL-LTSLFPALAEAGNTREDNFKHLGNDVKKRSEAGALQLOEVDQ 65
 OY 66 NPL-DEHEPELY-----TTALENKTMLINCAL-----NODIMRLACYDTLVHGE--- 109
 Db 66 MTLVDQHGKIDLRGMSTHGLQWPEIILNDNAYKALANDMDSNMIRLAMY---VGENY 121
 OY 110 --TPAVIKTK-----RSIRIDETI--WQT-----IKGRP 134
 Db 122 ATNPelikORVIDGLEIENDMYIVDWHVHAPGDPDPVYAGAKDFREIRALYPPNP 181
 OY 135 QVYQETTPD-----IFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDDLRN---N 183
 Db 182 HTIYELANEPSSNNNGAGIPNNEEGMAVK-----ETADPIVEMLRSGNADNDNIIIVG 236
 OY 184 TPLMSSRP-----HNPMYVLPFIMHGKPNRSPTPSHEARQFTPNERRAPELKR-- 231
 Db 237 SPNWSGRDPLADNPIDIDHHTMYTVHFTGSHAASTESTYSPF---TPNSEKGNVMSNT 291
 OY 232 --FQVSVKKAADLMGT-----DSDLMFGYTOO-----SHWQIFNGKNS-- 269
 Db 292 RYALENGVAFATE--WGTQASGDGPGYFDEADVWIEPLNENNISMANNSLTN-KNEVS 348
 OY 270 ---RFRV-----HDYQF-EIFLP-----QPV---YSDLPW-- 293
 Db 349 GATTFPELGKSNATNLDPGPDQVWPEELSLSGEYVRAIKGVNTEPDRKTYTVLWDF 408
 OY 294 -DGKVRMIGCAVHHSNGESAKLSRSMRAYLMAEM-----KNLTWPRIMGR 342
 Db 409 NDGTRQ--GFGVSDSPKKEILAYVDNENNTLKVSGOLDVSDNVDGNFANARLSANGCK 466

RESULT 15 T20122

hypochemical protein F25B3.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T20122; T21324
 R:McMurray, A.
 submitted to the EMBL data library, April 1996
 A:Reference number: Z19225
 A:Accession: T20122

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W1L>
A:Cross-references: EMBL:270750; PIDN:CAA94750.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone C50F4
R.Gardner, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19406
A:Accession: T21324
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W12>
A:Cross-references: EMBL:270752; PIDN:CAA94758.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone F25B3
C:Genetics:
A:Gene: CESP:F25B3.1
A:Map position: 5
A:Introns: 25/3; 47/2; 199/1; 242/1; 315/2; 544/3; 600/1; 626/1; 646/1; 666/3; 706/3; 75

Query Match 4.5%; Score 105.5; DB 2; Length 901;

Best Local Similarity 19.6%; Pred. No. 2.8;

Matches 74; Conservative 60; Mismatches 174; Indels 69; Gaps 15;

QY 31 VAVVDEVRSENDLGO-----DNELPIDVQASATQASATOTANPLDEHEPELYTTALENK 83
DB 468 VAMITEIRNOKDLEAVDVHRIPEVPTPOLASRNPALNOPTDADDDAETSNNRFFRS 527
QY 84 TMLI-----NCSALNODIMRLACYDTLVHGETPAVITKRSIRLDETIWQT 129
DB 528 NVSITWTPGVGAIIRASNRASPSKRDELROARDLIEKSTTPATAPNSRKASDEERRRE 587
QY 130 I-----KGPQVYVQETTPITFLMGNEKMLTKKAKOLEYAKOF-----TPLSLSD 178
DB 588 VRRLNEKQHPRTAIPSTSSSPPTFRRIDGSMT--DLRIELDVHKKRKPSPPTLVKQ 645
QY 179 LDRNNTPLMSRPHNPVYLPITFMHGKPNRSPN--TPSHEA--ROFTPNEFPAPELKF 232
DB 646 YDPNDF-----PH-----VPAIGRGRTNGNRDSSTPSSASTFDKRYKRGSMRSALKE 694
QY 233 QVSVKVK---AAEDLMGTDSDLMFGYTQO--SHWQ---IFNGKNSRPPRVHDYQPEIF 282
DB 695 SLQLMAKQYGMGNDPESSQDALATPTKFKFSQWEEKVDVDEGTANELVRIDERISDIT 754
QY 283 LTGPVYSDLPMWDKVRIGGAVHSHNGESAKLSRSNNRAYLMAGMEMKN-LIYMPRIWG 341
DB 755 AQADVITD-----KIRETEVGSSEEMLTASYLETNERNTLVHROEYVNIITIRQVTS 809
QY 342 RIFKEGS--GSOPPDNP 356
DB 810 EIDQLGKQINEVPDDFP 826

Search completed: November 30, 2001, 14:18:35
Job time: 257 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:27:03 ; Search time 28.02 Seconds

(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-6
Perfect score: 2360
Sequence: 1 MKVSLSTLTLSLSCFAILA.....YNHEATSPGVGLMLNDMMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	15.8	286	1	PAL_KLEPN
2	358	15.2	289	1	PAL_SALTY
3	357	15.1	289	1	PAL_ECOLI
4	321.5	13.6	289	1	PAL_PROVU
5	109.5	4.6	800	1	GMD_BACSI
6	98.5	4.2	810	1	AMD_YEAST
7	97.5	4.1	1478	1	BCK1_YEAST
8	95	4.0	719	1	PBPA_STRPN
9	94.5	4.0	397	1	GUN_PABPO
10	94	4.0	716	1	BAC2_MOUSE
11	94	4.0	1115	1	DP3A_BACSU
12	92.5	3.9	849	1	SRK6_BRAOL
13	92.5	3.9	1379	1	MET_MOUSE
14	92	3.9	324	1	VL70_ARCFU
15	92	3.9	888	1	VJH0_YEAST
16	91.5	3.9	486	1	VG14_BPM15
17	91.5	3.9	969	1	SACB_STRSL
18	91	3.9	496	1	CATP_DICDI
19	90.5	3.8	467	1	INVO_MOUSE
20	90.5	3.8	1024	1	Y075_MYCGE
21	90	3.8	1158	1	R114_HUMAN
22	90	3.8	1788	1	VP72_CAEEL
23	89	3.8	521	1	NPRE_BACAM
24	88.5	3.8	825	1	GUN3_BACSI
25	88.5	3.8	1350	1	MET_HUMAN
26	88	3.7	483	1	MURE_CHLTR
27	88	3.7	537	1	PAH2_MOUSE
28	88	3.7	1382	1	MET_RAT
29	87.5	3.7	478	1	DHGB_ACTICA
30	87.5	3.7	503	1	CD44_RAT
31	87.5	3.7	669	1	AMY_ALTHA
32	87	3.7	637	1	PBPA_STROR
33	87	3.7	666	1	PDI4_RAT

34	87	3.7	828	1	BGAL_BRAOL	P49676	brassica o1
35	87	3.7	842	1	LPFC_SALTY	P43662	salmonella
36	87	3.7	913	1	YGJ3_YEAST	P53148	saccharomyc
37	87	3.7	1474	1	A2MG_HUMAN	P01023	homo sapien
38	86	3.6	1008	1	SN14_YEAST	P36048	saccharomyc
39	86	3.6	1087	1	XYNX_CLODM	P38535	clostridium
40	85.5	3.6	353	1	DCUP_BACSU	P32395	bacillus su
41	85.5	3.6	790	1	SEIL_MOUSE	O92246	mus musculu
42	85.5	3.6	794	1	SEIL_HUMAN	O9ub22	homo sapien
43	85.5	3.6	1374	1	KC9A_SCHRO	O09884	schizosach
44	85	3.6	491	1	TY3H_PNASP	P11982	phasiandae
45	85	3.6	525	1	MP11_RAT	P48965	rattus norv

ALIGNMENTS

RESULT ID	PAL_KLEPN	STANDARD	PRT	286 AA.
AC	P37446:			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
CN	PLDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
OX	NCBI_Taxid=573;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94131966; PubMed=8300539;			
RA	Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.;			
RT	"Molecular characterization of enterobacterial plda genes encoding			
RT	outer membrane phospholipase A.";			
RT	J. Bacteriol. 176:861-870(1994)			
RL	J. Bacteriol. 176:861-870(1994)			
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	-1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X76901; CA54223.1; -.			
DR	PIR: B36971; B36971.			
DR	PIR: S40129; S40129.			
DR	InterPro: IPR003187; PLA1.			
DR	Pfam: PF02253; PLA1; 1.			
KW	Hydrolase; lipid degradation; Outer membrane; Signal; Calcium.			
FT	SIGNAL	1	20	BY SIMILARITY.
FT	CHAIN	21	286	PHOSPHOLIPASE A1.
FT	ACT SITE	161	161	BY SIMILARITY.
SO	SEQUENCE	286 AA;	32544 MW;	3E39F863085108A3 CRC64;

Query Match 15.8%; Score 373; DB 1; Length 286;
Best local Similarity 42.1%; Pred. No. 5.0e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

[illegible]

```

Best Local Similarity      39.3%, Pref. No. 9.2e-22;
Matches      84; Conservative      30; Mismatches      90; Indels      10; Gaps      4;

QY      226  RAPELKFQVSVKVAKAEDLMGTDSDLMVGYTQOOSHQMFJFNKNSRPFVRHDPYQPEIFLTQ 285
      1-|-|-|:-:      :      :      :      :      :      :      :      :      :      :
Db      83  RKDEVKFLQSLIAFLPMRGILGPNISVLAGSYTQKSMQWQJNSKSSPFPRETYEPQLGLGF 142

QY      266  PV-YSDLEPMDGKVRNIGGAVHHNSGSEAKLSRSRNRAYLMAAGHEKMLTYMPRIWGRIF 344
      :      :      :      :      :      :      :      :      :      :      :      :
Db      143  AADYFPAG--TLRHEVGGYVHNSGRSDPTSRFTWNRRLTYRLMAENGWMLVEVPRWYI- 199

QY      345  KEGSSGQPDNDPDLIDYGYGVDFEFLYOLEKNSISGVRFRNPPSGALDLDVYPLGK 404
      |||||      |      |      |      |      |      |      |      |      |      |
Db      200  -----GSDNDNPDLIRKYGAYQKLGIVL-GEAVLSAGQVNNMTYTGGAARVGLSTYPTK 253

QY      405  GISGFQIFQGTGOSLIDYNHEATSPFGVGLMLND 438
      :      :      :      :      :      :      :      :      :      :      :
Db      254  HVRLTYQVYSGGESLIDYNFNQTRFVGVMIND 287

RESULT      3
PAL_ECOLI
ID      PAL_ECOLI      STANDARD;      PRT;      289 AA.
AC      P00631;
AD      21-JUL-1986 (Rel. 01, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DR      20-AUG-2001 (Rel. 40, Last annotation update)
DE      PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DEREGENT-RESISTANT
DE      PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE      ACILHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN      PLA OR B3821 OR Z5342 OR ECS4751.
OS      Escherichia coli, and
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562, 83334;
RN      [1]
RN      SEQUENCE FROM N.A.
RA      MEDLINE=85157492; PubMed=6397464;
RA      Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,
RA      Inoue K., Ikeda H., Sekiguchi M., Nojima S.;
RA      "The DNA sequence encoding plidA gene, the structural gene for
RA      detergent-resistant phospholipase A of E. coli.";
RA      J. Biochem. 96:1655-1664(1984).
RL      [2]
RL      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RC      MEDLINE=92358234; PubMed=1379743;
RA      Daniels D.L., Plunkett G. IIf, Burland V.D., Blattner F.R.;
RA      "Analysis of the Escherichia coli genome: DNA sequence of the region
RA      from 84.5 to 86.5 minutes.";
RA      Science 257:771-778(1992).
RL      [3]
RL      REVISION TO 14-15.
RP      STRAIN=K12 / MG1655;
RC      MEDLINE=97426617; PubMed=92785903;
RA      Blattner F.R., Plunkett G. IIf, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RA      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [4]
RN      SEQUENCE FROM N.A.
RC      STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. IIf, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Postal G., Heckett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grothbeck E.J., Davis N.W., Lim A., Dolananta E.T., Potamousis K.,
RA      Apodaca E., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blattner F.R.;
RA      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RT

```

RL Nature 409:529-533(2001).
 [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RA MEDLINE-01576331; PubMed-1258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kihara S., Shiba T., Hattori M., Shingawa H., Escherichia coli
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP SEQUENCE OF 30-289 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-85003590; PubMed-6383820;
 RA de Geus P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
 RT "The pro- and mature forms of the E. coli K-12 outer membrane
 RT phospholipase A are identical.";
 RL EMBO J. 3:1799-1802(1984).
 RN [7]
 RP SEQUENCE OF 174-289 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-87115164; PubMed-3027506;
 RA Irino N., Nakayama K., Nakayama H.;
 RT "The recO gene of Escherichia coli K12: primary structure and
 RT evidence for SOS regulation.";
 RL Mol. Gen. Genet. 205:298-304(1986).
 RN [8]
 RP MUTAGENESIS OF SER-172.
 RX MEDLINE-94131966; PubMed-8300539;
 RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
 RT Verheij H.M., Tomassen J.;
 RT "Molecular characterization of enterobacterial pldA genes encoding
 RT outer membrane phospholipase A.";
 RL J. Bacteriol. 176:861-870(1994).
 RN [9]
 RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
 RX MEDLINE-91249806; PubMed-2040286;
 RA Horrevorts A.J.G., Verheij H.M., de Haas G.H.;
 RT "Inactivation of Escherichia coli outer-membrane phospholipase A by
 RT the affinity label hexadecanesulfonyl fluoride. Evidence for an
 RT active-site serine.";
 RL Eur. J. Biochem. 198:247-253(1991).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
 CC DOMINANT IN NORMAL GROWING CELLS.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
 CC LYSIS OR TEMPERATURE SHOCK.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X02143; CAA26081.1; -
 CC EMBL: M87049; AAA67617.1; -
 CC EMBL: AE000458; AAC76824.1; -
 CC EMBL: AE005613; AAG59017.1; -
 CC EMBL: AP002567; BAB38174.1; -
 CC EMBL: M30198; AAA24516.1; -
 CC PIR: A00771; PSECA.

DR PIR: A22133; PSECA.
 DR PIR: S30711; S30711.
 DR Ecogene: EG10738; PldA.
 DR InterPro: IPR003187; PldA.
 DR Pfam: PF02253; PldA; 1.
 KW Hydrolysis; Lipid degradation; Outer membrane; Signal; Calcium;
 KW Complete proteome.
 FT SIGNAL 1 20
 FT CHAIN 21 289 PHOSPHOLIPASE A1.
 FT ACT_SITE 164 164
 FT MOTIF 172 172 S->F: INACTIVE PROTEIN.
 FT CONFLICT 14 15 LP -> FA (IN REF. 2).
 FT CONFLICT 30 33 DAPA -> MTRQ (IN REF. 6).
 SQ SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CRC64;
 Query Match 15.1%; Score 357; DB 1; Length 289;
 Best Local Similarity 39.3%; Pred. No. 1.1e-21;
 Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;
 QY 226 RAPELKQVSVKVAEDLMQTSDFYTGQSHWQIFNKNRRPRVHYDFEFLIQ 285
 DB 83 RKDEVKFQLSLAPFLMGLGPNVLSGASTYQKSMQLNSSESSPRETYEPQLFLGF 142
 QY 286 PV-YSDLPMDGKVMIGAVHSHNGESAKLSRSMNRAYLAGEMKNLTVMPRMGRIF 344
 DB 143 ATDYRFAGW--TLNDVEMKYNHDSNGSDPTSRSMNRLYTLMAENGWLVKVPWTVV- 199
 QY 345 KEGSGSQDDNPDLIDYGYGVDFLYQLENKSNISGTVRYNPRSGKALQLDYVPLGK 404
 DB 200 -----GMTDDNPDTIKMGVYQLKIGYHL-GDAVLSAKGQNMWTVYGGALGLSTPLTK 253
 QY 405 GISGFQIFQCYGOSLDYHNHATSFVGLMLND 438
 DB 254 HVRLYTVSGYSGESLIDYFNQTRVGVMLND 287
 RESULT 4
 ID PAL_PROVU STANDARD; PRT; 289 AA.
 AC P37447;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, last annotation update)
 DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
 DE PHOSPHOLIPASE A) (OR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
 DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
 GN PldA.
 OS Proteus vulgaris.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Proteus.
 OC NCBI_TaxID=585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94131966; PubMed-8300539;
 RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
 RT Verheij H.M., Tomassen J.;
 RT "Molecular characterization of enterobacterial pldA genes encoding
 RT outer membrane phospholipase A.";
 RL J. Bacteriol. 176:861-870(1994).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X76902; CAA54224.1; -
DR PIR: C36971; C36971.
DR PIR: S40130; S40130.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 1 21 PHOSPHOLIPASE A1.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 289 AA; 32944 MW; D75516CFB406997 CRC64;

Query Match 13.6%; Score 321.5; DB 1; Length 289;
Best Local Similarity 33.3%; Pred. No. 8.3e-19;
Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

QY 179 LDRNPTLWSSRPHNPVYLPIFMHGPNSRSPNTPSHEARQTPNEFRAPDLKQVSVKV 238
DB 43 LOEHNP-FTLPYESPNTLYTY-----TSDLNKKAIESYNSDNA-NKDEYKFLSLAF 95
QY 239 KAEDLMGTDSDLFQYQOOSHWOIFNGKNSRPFVHYDIOPEIFLTPQV-YSDLPWDGKV 297
DB 96 PLMRGILGNSLIGASYTORSMWOLSTNGESAPFRETYEPOLFGLFADYSGVW--TL 153
QY 298 RMIGGAVHNSGESAKLSRSNRAVILMAGMEKMLTVPRIRMGRIKFGSSQSDPDND 357
DB 154 KAEFGYHNSGSRDPNRSNNRLXSLMAONGMLVEKFWYI-----GDTSDNKN 207
QY 358 ILDYGYGDVRLYOLENKSNIQTVRYNPRSGKALQLDYVPLGKISGYFOIYOGYG 417
DB 208 IRYMGYYQLKIGYOL-GEAVLSAKGYVMNTGYGGAELGVSYPITKHRYFTQYYSYG 266
QY 418 QSLIDYNNHATSFYGLMLND 438
DB 267 ESLIDYDFNQTRVGMGMVLMND 287

RESULT 5
GUN_BACSI
ID GUN_BACSI STANDARD: PRT: 800 AA.
AC P06564;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE (ALKALINE CELLULOSE).
OS Bacillus sp. (strain 1139).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1411;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=87085443; PubMed=3098909;
RA Fukumori F., Kudo T., Narahashi Y., Horikoshi K.;
RT "Molecular cloning and nucleotide sequence of the alkaline cellulase
RT gene from the alkalophilic Bacillus sp. strain 1139.";
RU J. Gen. Microbiol. 132:2329-2335(1986).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1- MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES
CC ONE ALKALINE CELLULOSE (PH OPTIMUM 9.0). THIS STRAIN IS NOT A TRUE
CC CELLULOTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO
CC HYDROLYSE NATIVE CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC ."

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M15743; AAA22305.1; -
DR PIR: D00066; BAA00045.1; -
DR PIR: A29003; A29003.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF001050; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 1 31 800 ENDOGLUCANASE.
FT ACT_SITE 190 190 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 305 305 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 800 AA; 88602 MW; 7CCA4D7B6DAD55CF CRC64;

Query Match 4.6%; Score 109.5; DB 1; Length 800;
Best Local Similarity 19.8%; Pred. No. 0.47;
Matches 95; Conservative 54; Mismatches 163; Indels 169; Gaps 25;

QY 9 TSLISCFALIAIQAKAVNVAATVDEYRSEND--LGDNELPIDVQASATQASATDTA 65
DB 7 TKLLSSILILVL-LLSLFTALAEAGNTREDNFKHLGLNDVYKRSEAGALQLEVDQ 65
QY 66 NPL-DEHEPELY-----TFALENKTMILNCSAL-----NODIMRLACY-DTLVHGEP 111
DB 66 MTLVQHGKIDLRGKSTHGLQWPEIILNDNAYKALANDMESMIRLAMYDENGYASNP 125
QY 112 AVIKTKRSIRIDEIT-----NOT-----IKGRQVYV 138
DB 126 ELIKSRVIGIDILAIENDMYVIVDWHVHAPGDPDVPVYAGADEFFRDIALALYPNPHIY 185
QY 139 QETDPR-----IFLMGKEKMLTKKAKOLEYAAKOPTSLSFIDLNR-----NTPLW 187
DB 186 ELANEPSSNNNGACGIPNNEEGMAVKK-----EYADPIYEMLRDGNADNDIITIGSPW 240
QY 188 SSRP-----HNPNVLPPIF--MHGKPNRS--PNTPSHEARQTPNEFRAPDLKQ 233
DB 241 SQRPDLADNPIDDHHTYTVHFGYSHASRESYRPETPNBERGNVMSNTYTA-----LE 296
QY 234 VSVKVKAAEDLMGT-----DSDLMFGYIQO-----SHWOIFNGKNS-----RP 271
DB 297 NCVAVAFATE--WGTSQANGDGPRFDEADVWIEFLNENNISWANSLSLV-KNEVSGAFRP 353
QY 272 FVHV-----DYOEIFLTPQVYSDLPW---D 294
DB 354 FELGKSNATSLDPRDQWVPEELSLGEXYVARIKGVYEP--IDRKYYKVLMDEND 410
QY 295 GKVRMIGGAVHNSGES-----AKLSRSNRAVILMAGMEKMLTVPRIRWG 341
DB 411 GTRKGQGV-----NGDSVEDVYIENEGALKLSGLDASNDVSGENYANRLASDNG 464
QY 342 R 342
DB 465 K 465

RESULT 6
AAMD_YEAST
ID AAMD_YEAST STANDARD: PRT: 810 AA.
AC P15274;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE AMP DEAMINASE (EC 3.5.4.6) (MYOADENYLATE DEAMINASE).
GN AMD OR AMD OR YML035C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90105403; PubMed-2690949;
 RA Meyer S.L., Kvalnes-Rick K.L., Schramm V.L.;
 RT "Characterization of AMD, the AMP deaminase gene in yeast. Production
 of and strain, cloning, nucleotide sequence, and properties of the
 RT protein.";
 RL Biochemistry 28:8734-8743(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-5288C / AB972;
 RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
 CC METABOLISM.
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M30449; AAA34420.1; -
 CC EMBL: Z46659; CAA86620.1; -
 CC PIR: A33365; A33365.
 CC SGD: S0004498; AMD1.
 CC InterPro: IPR001365; A_deaminase.
 CC Pfam: PF00962; A_deaminase; 1.
 CC PROSITE: PS00485; A_DEAMINASE; 1.
 CC HydroLase: Nucleotide metabolism.
 FT ACT_SITE 422 422
 FT ACT_SITE 631 631 POTENTIAL.
 FT ACT_SITE 707 707 POTENTIAL.
 FT ACT_SITE 708 708 POTENTIAL.
 FT ACT_SITE 568 568 F -> C (IN REF. 1).
 FT CONFLICT 568 568
 FT SEQUENCE 810 AA; 93301 MW; 7A6DCB43B9B45C93 CRC64;
 SQ

Query Match 4.28; Score 98.5; DB 1; Length 810;
 Best Local Similarity 21.38; Pred. No. 3.7;
 Matches 74; Conservative 39; Mismatches 132; Indels 103; Gaps 15;

QY 46 DNELPIDVOSATOSASNDTAN-----PLDEHEPELXTTALLENKTM 86
 DB 48 DEPRPLEQDSHESLADSNANFSTYENQOQLBENGTKALDBHDS--HSALEOPSHS 105
 QY 87 INCSALNODIMRLACYDTLVH-----GETPAVITKTRSRILDETITQTKGPOVVYQET 141
 DB 106 TNCSSSIAMMNG-HDSADHASNGSGKPRITLSASQHLIPETL-KSFAGAPVYNNQVR 163
 QY 142 TDPIFLMG-----NEKGLTKKAKOLEVAKOFTPLSLSFDDRRNTP 186
 DB 164 TSASYKMGMLADDAOQFLDDPSSSELIDLYSKVAECNLRKAKYQTTISVQDDNPFKPKPG 223
 QY 187 W--SSRPHNMY-----VLPFMHGKPNRSPMPRSHARQFPNFRRAELKQVSVKV 238
 DB 224 WVVYPPRPKSYSDTKTVV-----TNKPRDAEYDFTKCE----- 260
 QY 239 KAAEDLWGTDSIDMFQYTOOSHQIFENGKNSRFRVHDYQPEIFLTQPVYSDLPWDGKVR 298
 DB 261 -----IPGEEDPDMETLNDSDYV-----HRSKGTDELLAQIPTLDYLDLE-----K 305
 QY 299 MIGHGAVHNSGSAK-----LSRSNRAVILMACMEKKNLTPMR 338
 DB 306 MISIS-----SDGPAKSPAYRRLOYLEARNMLYYLL--NEYQETSVSRR 347

RESULT 7
 BCK1_YEAST
 ID BCK1_YEAST STANDARD; PRT; 1478 AA.
 AC Q01389; P32894;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SERINE/THREONINE PROTEIN KINASE BCK1/SLK1/SSP31 (EC 2.7.-.-).
 GN BCK1 OR SLK1 OR SSP31 OR LAS3 OR YJ095W OR J0906.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-92186847; PubMed-1545797;
 RA Costigan C., Gehring S., Snyder M.;
 RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
 RT homolog implicated in yeast cell morphogenesis and cell growth.";
 RL Mol. Cell. Biol. 12:1162-1178(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-92104496; PubMed-1840547;
 RA Irie K., Araki H., Oshima Y.;
 RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
 RT involved in plasmid maintenance in Saccharomyces cerevisiae.";
 RL Gene 108:139-144(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BG123;
 RX MEDLINE-92107166; PubMed-1729597;
 RA Lee K.S., Levin D.E.;
 RT "Dominant mutations in a gene encoding a putative protein kinase
 RT (BCK1) bypass the requirement for a Saccharomyces cerevisiae protein
 RT kinase C homolog.";
 RL Mol. Cell. Biol. 12:172-182(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-5288C;
 RX MEDLINE-95176706; PubMed-7871887;
 RA Miosga T., Boles E., Schaaff-Gerstenschlaeger I., Schmitt S.,
 RA Zimmermann F.K.;
 RT "Sequence and function analysis of a 9.74 kb fragment of
 RT Saccharomyces cerevisiae chromosome X including the BCK1 gene.";
 RL Yeast 10:1461-1488(1994).
 RN [5]
 RP SEQUENCE OF 602-1104 FROM N.A.
 RA Cusick M.E.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
 CC TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
 CC AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
 CC INVOLVE THE KINASE PKC1 THAT MAY ACT ON THIS KINASE. BCK1 PROBABLY
 CC PHOSPHORYLATES MKR1 AND MKR2 WHICH THEMSELVES PHOSPHORYLATE THE
 CC MPK1 KINASE.
 CC -1- SUBCELLULAR LOCATION: CYTOSOL; PLASMA MEMBRANE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M84389; -; NOT_ANNOTATED_CDS.
 CC EMBL: D10389; BAA01226.1; -;
 CC EMBL: X60227; CAA42788.1; -;
 CC EMBL: X77923; CAA54896.1; -;
 CC EMBL: Z49370; CAA89389.1; -;

```

DR EMBL; Z49369; CAA89388.1; -
DR EMBL; M88604; AAA21179.1; -
DR PIR; S20117; S20117.
DR PIR; S22285; S22285.
DR PIR; J01118; J01118.
DR PIR; J01432; J01432.
DR HSSP; P24941; 1A01.
DR SGD; S0003631; BCK1.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kin_atacste.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR TRANSFERASE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
ATP-binding; Phosphorylation.
FT DOMAIN 1175 1440 PROTEIN KINASE.
FT NP_BIND 1181 1189 ATP (BY SIMILARITY).
FT BINDING 1204 1204 ATP (BY SIMILARITY).
FT ACT_SITE 1303 1303 BY SIMILARITY.
FT MOD_RES 1134 1134 PHOSPHORYLATION (BY PKC) (POTENTIAL).
FT VARIANT 1119 1119 T -> P (IN BCK1-19; ACTIVATION).
FT VARIANT 1120 1120 I -> K (IN BCK1-11; ACTIVATION).
FT VARIANT 1120 1120 G -> V (IN BCK1-16; ACTIVATION).
FT VARIANT 1146 1146 A -> P (IN BCK1-20; ACTIVATION).
FT VARIANT 1174 1174 A -> I (IN REF. 2).
FT CONFLICT 59 59 E -> I (IN REF. 3).
FT CONFLICT 79 79 F -> V (IN REF. 3).
FT CONFLICT 264 264 A -> P (IN REF. 3).
FT CONFLICT 279 279 N -> I (IN REF. 3).
FT CONFLICT 703 714 RYPTQPSIYDR -> STPKRIVTMT (IN REF. 3).
FT CONFLICT 795 795 S -> A (IN REF. 3).
FT CONFLICT 802 802 L -> V (IN REF. 3).
FT CONFLICT 808 808 A -> S (IN REF. 3).
FT CONFLICT 903 903 T -> N (IN REF. 3).
FT CONFLICT 919 919 T -> N (IN REF. 3).
FT CONFLICT 960 962 ADA -> RDR (IN REF. 5).
FT CONFLICT 1086 1104 RPVPDSSEYEFIDGLNCK -> VPIAHTSSYRMDLTVKIN
H (IN REF. 5).
SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5BB3 CRC64;

```

```

Query Match 4.1%; Score 97.5; DB 1; Length 1478;
Best Local Similarity 23.0%; Pred. No. 10;
Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;

```

```

OY 2 KVSISTFLS-----ILSCFALIAIOAKAVNPVAFVDEVSENDLQDNELPI---DV 53
DB 912 KVNNSNSTVTSNSLFTYSPSPILKRGNSKRYVSSISAAD-IFEENDITFADAPPFDSDD 970
OY 54 OSATOSASTDTANPLDEHEPELYTTALENK--TMLINCSALNODIMRLACYDTLVHGETP 111
DB 971 SDDSSSSDDDIIMSKKKAPE---TNNENKKDEKSDNSTSHDELF-----QTO 1019
OY 112 AVIKRKSRIRLDEITWQITKGGPOVYVQETDPIFLMGNEKGMILTKKAKOLEVAAKQFT 171
DB 1020 DKMERK-----MFRPSPEVAVYO-NLEKFRANLDKPIR-----EGIASPTS 1061
OY 172 PLISLFDLDRNNT-----PLWSSRPHPMYLPIFMHG-----KPNSPNTP----- 213
DB 1062 PKSLDLSLSPKNVASSRKEPSTPSRPVPPDSSYEFTIQGLGKNKRPLOAKTPKTKTIR 1121
OY 214 --SHEARQFTNEFRAPRLKQVSVKVAE-DLWGTSDMLFGYTOOSHQIFNGKNSR 270
DB 1122 TIAHNASLARKN-----SVKLKRLQNTKMMGT---RMVEVTEHNHVSINKAKSK 1167
OY 271 PFRVHDYPELFLQPVYSDLPMDGKVMIGGA-----VHNSNGSAKLSR----- 317
DB 1168 -----GEYKEFEAW-MKGEMIKGSGVAYVLLANVTGEMAVKQVEVPKY 1211
OY 318 -SWNRAYLMAGEMKKNLVMPRIMGRIRKEGSGSPDNPILDVYGGVDRFLQLQLENK 376

```

```

DB 1212 SQONFALIS-----TV-----EALRSEVSTLKDLHLNIVQYLG-----ENK 1249
OY 377 SNISGTVARNPRSGKALQLDVYPLGKIGISGYFOIFOGYGSLLIDYNEATSPGVGL 434
DB 1250 NNIV-----SLFLEYV--AGGSVGSLLIMYGRFDEPLI--KRLTLYQVAKGL 1291

RESULT 8
ID PBPB_STRPN STANDARD; PRT; 719 AA.
AC 004707;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (PBP-1A) (EXPORTED PROTEIN 2).
GN PONA OR EXP2.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=45607, AND 63915;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain."
RL EMBO J. 11:3831-3836(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Briese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins
RT from Streptococcus pneumoniae and Streptococcus oralis with high
RT homology to Escherichia coli penicillin-binding proteins 1a and 1b."
RL J. Bacteriol. 174:4517-4523(1992).
RN [3]
RP SEQUENCE OF 293-369 FROM N.A.
RC STRAIN=RX6;
RX MEDLINE=95020625; PubMed=7934910;
RA Pearce B.J., Yin Y.B., Masure H.R.;
RT "Genetic identification of exported proteins in Streptococcus
RT pneumoniae."
RL Mol. Microbiol. 9:1037-1050(1993).
CC -1- FUNCTION: CELL WALL FORMATION.
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL; X67873; CAA48073.1; -
DB EMBL; X67872; CAA48072.1; -
DB EMBL; M90527; AAA26956.1; -
DB PIR; S28038; S28038.
DB InterPro; IPR001264; Transglycosyl.
DB InterPro; IPR001460; Transpeptidase.
DB Pfam; PF00912; Transglycosyl; 1.
DB Pfam; PF00905; Transpeptidase; 1.
DB PRODOM; PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme.
FT ACT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT DOMAIN 657 683 SER-RICH.
FT NP_BIND 124 124 T -> A (IN STRAIN R6).
FT VARIANT 386 386 V -> I (IN STRAIN 63915).

```

FT VARIANT 388 388 D -> E (IN STRAIN R6).
 FT VARIANT 397 397 E -> K (IN STRAIN 63915).
 FT VARIANT 523 523 M -> I (IN STRAIN 63915).
 FT VARIANT 533 533 E -> D (IN STRAINS 63915 AND R6).
 FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
 FT VARIANT 657 657 S -> N (IN STRAIN 63915).
 SQ SEQUENCE 719 AA: 79745 MW: 58C36A93BFA970A CRC64;

Query Match 4.0%; Score 95; DB 1; Length 719;
 Best Local Similarity 20.1%; Pred. No. 6;
 Matches 96; Conservative 70; Mismatches 196; Indels 116; Gaps 25;

QY 3 VLSITLSTLSTL-----LSCFAILLAIQAKAVPRPVAVDRENDL 43
 DB 13 LSTISLSTLVAIAVLGCGVFYVYVSKAPSLSEKSLVATTSKIKYDKKNOLLADLGGSERRV 72
 QY 44 -GQDNELPDIQVQATQS-----ASTDTANPLDEHEBELYTTALENKTMLINCSALN 93
 DB 73 MDAQNDIPFDLVKAIYSIEDHREFDHGIDTIRILGAFILNLSQNSLQGG-----STLT 126
 QY 94 QDIIRLACDTLVHGTGTPAVIKTKRSIRIDETIWTQTKGPOVVOET---TDPIFLMG 149
 DB 127 QOLIKLTFTSTSDQTI-----RKAQEAAILQLEQKATKQEIILTYINKYMSN 178
 QY 150 NEKGMLTK-----KAKOLE-----YAAKQTFPLS--LSFDDLRNNTPLMSSRP 191
 DB 179 GNYGMQTAONNYKQDLNNLSLPQALLAGMPQAPNOYDPSHPREAQDRNLVL--SEM 236
 QY 192 HNPVYLPFLFMHGKPRSPNTPSHEARQFTPNFEFRAPE-----LKFQVS--VKAKAEDLW 245
 DB 237 KNGGYI-----SAEQYKAVNTPITDGLQSLKASNSNPAYMDNLKEVINVEETGYNLL 292
 QY 246 GTDSDLMEGTQ--QSH--NOIENGKNSRPRVHDYQPELFLQPVYSDLPWDGKVRMIGM 302
 DB 293 TTGMDEVYTVNDQAKHLDIYNTDEVYVAPDDELQ-----VASTIVDVS--NGKV--IAQL 345
 QY 303 GAVVHNSN-----GESAKLSRSNNRAYLMAGMEMKNLT--VMPRI-----WRIKREGS 348
 DB 346 GARHOSSNVSFGINQAVETNRDW-----GSTMKPRTDYAPALEGYVDSTATYIDEP 398
 QY 349 GSGQPDNDPILDY--YGY--GDVRFYQLEKNSNISGTVRYN-----PRSGKALQLDY 398
 DB 399 YNPGNTFVYNNMDRGYFGNITLQYALQSSRNVPVETLNKVKLNKAKFLNGLGIDY 456

RESULT 9
 GUN_PAPEO STANDARD; PRT; 397 AA.
 ID GUN_PAPEO P23548;
 AC 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ENDOGUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULOSE).
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 OC Bacacteria: Firmicutes: Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Paenibacillus.
 OK NCBI_TaxID=1406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90170877; PubMed=2307659;
 RA Baird S.D., Johnson D.A., Seligy V.L.;
 RT "Molecular cloning, expression, and characterization of
 endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
 circulans";
 RL J. Bacteriol. 172:1576-1586(1990).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
 CC LINKAGES IN CELLULOSE.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
 CC HYDROLASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M33791; AAA2631.1; -
 DR PIR: A35136; A35136.
 DR HSSP: P54583; 1ECE.
 DR Interpro: IPR001547; Glyco_hydro_f5.
 DR Pfam: PF00150; cellulase; 1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
 KW Cellulose degradation: Hydrolase; Glycosidase.
 FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
 SQ SEQUENCE 397 AA: 44357 MW: B9C2E802C04F0A2A CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 397;
 Best Local Similarity 25.7%; Pred. No. 2.9;
 Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY 301 GMGAVHNSGSAKLSRSNNRAYLMAGMEMKNLTVMPIRIGR-----IFKEG----- 347
 DB 43 GNKIVDESGKEAFLNGLNW-----FGLETPRNT--LHGIMSRSMDDMLDQVKKEGYNLIR 95
 QY 348 -----SGSQPD-----NDIILDYGYGDVRFYQLEKNSNISGTY-----RINPRS 389
 DB 96 LPYSNQLFDSSSRPSIDYKKNPDLV---GLNPDIQIMDKLEKAGORGIOIILDRHRPGS 152
 QY 390 GRGALQLDYV--YPLGIGISGYQIFQGY-----GQSLIDVNHETSRYGVLMDW 439
 DB 153 G-GQSELMWTSQYIPESRWISDMKMLADRYKNNPTVIGADLHNPFGQASMGTONASTDW 210

RESULT 10
 BAC2_MOUSE
 ID BAC2_MOUSE P97303;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTF AND CNC HOMOLOG 2).
 GN BACH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=97042348; PubMed=8887638;
 RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
 RA Yamamoto M., Igarashi K.;
 RT "Bach proteins belong to a novel family of BTF-basic leucine zipper
 RT transcription factors that interact with Max and regulate
 RT transcription through the NF-E2 site";
 RL Mol. Cell. Biol. 16:6083-6095(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
 CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
 CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
 CC MAX.
 CC -1- SUBUNIT: HETERODIMER OF BACH2 AND MAX.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
 CC NEURONAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D86604; BAA1338.1; -
 DR HSSP: P05412; IFOS.
 DR MGSD: MGI:894679; BacH2.
 DR InterPro: IPR000210; BTB_POZ.
 DR InterPro: IPR001871; bZIP.
 DR Pfam: PF00651; BTB; 1.
 DR Pfam: PF00170; bZIP; 1.
 DR SMART: SM00338; BRLZ; 1.
 DR SMART: SM00225; BTB; 1.
 DR PROSITE: PS50097; BTB; 1.
 DR PROSITE: PS00036; bZIP_BASIC; 1.
 KW Transcription regulation; Activator; Repressor; DNA-binding;
 KW Nuclear protein.
 FT DOMAIN 37 103 BTB.
 FT DNA_BIND 162 168 BASIC MOTIF.
 FT DOMAIN 527 542 BASIC MOTIF.
 FT DOMAIN 550 572 LEUCINE-ZIPPER.
 SQ SEQUENCE 716 AA; 78935 MW; 913283731AE24333 CRC64;

Query Match 4.0%; Score 94; DB 1; Length 716;
 Best Local Similarity 23.1%; Pred. No. 7.2; Mismatches 140; Indels 156; Gaps 24;
 Matches 101; Conservative 40;

OY 14 SCFALIAIQAKAVNPVAFV---DEV---RSENDG-----QDNELPIDVQSATQSAST 62
 DB 124 SCFSLQQLQNLNREDO--LFVCRKDSACORQEDHNSAGEEEETMDSTAMACAT 181
 OY 63 D-----TANPLDEH-----EPELYTALENKTMINCASALNQ---DIMRLAC 101
 DB 182 DQMLDPDISFEATAPVAKEKALLPESEVPTDKENSEK---GALIQYPRKKYQLAC 237
 OY 102 ---YDLVHG-----ETPAVTKTKRSTL---DET 125
 DB 238 TKNVSAHSFGSGFASFSSEDSPPGSLKPGCLPMQIKSEPSSEETESTLCLSGDET 297
 OY 126 IWQITKGRPOVYQETTPPI-----FLMGNEKMLTKK 158
 DB 298 ---DIKDRPGDVEMRKQSPARTPTRTGAACIDRSRSVSSPCLRLSFGITKGV--- 350
 OY 159 DAKOLEYAKOFTPLSLF-----DIDRNNPLMWS--RPHNPMVYLPFHMKG 205
 DB 351 ESTGIPSTSQ--PLVRSACPFNKGISQGLKDTYPLAGNYGQPHYGQKXVSFAMGS 408
 OY 206 PNRSP-----NTPSHANRQFTPNFRAPELK-----FQVSVKVAEAEDLMGTDSL 251
 DB 409 PLRGGPETLCFFSSPCSQARFLATEHQEPGLMGDGMYNQVPOIK--CEOSYSTNSSD 467
 OY 252 WEGYTQASHMOIFNGKNSRPFVHDYQPEIFLTQPV--YSDLPMGKRYMIGMGAHVHNS 309
 DB 468 ESG-----SEFSADSESCPVODRGQEVKLPFVQDITDPLRNDPQMIMK---HKLIT 516
 OY 310 GESAKL-----SRSMNR 321
 DB 517 SEQLEFIHDIRRRSKNR 533

RESULT 11
 DP3A_BACSU STANDARD; PRT; 1115 AA.
 AC 034623;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE DNA POLYMERASE III ALPHA SUBUNIT (EC 2.7.7.7).
 GN DNAE.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE:9804867; PubMed:9387221;
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 in the 200 kb rnb-dnaB region";
 RL Microbiology 143:3431-3441(1997).
 CC -I- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
 CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
 CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
 CC -I- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
 CC N PYROPHOSPHATE + DNA(N).
 CC -I- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
 CC CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
 CC DIMERIZATION TO FORM THE POLIIT' COMPLEX. POLIIT' ASSOCIATES WITH
 CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
 CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. DNAA
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF008220; AAC00338.1; -
 DR EMBL: Z99118; CAB1483.1; -
 DR Subtilist; BG12583; dnaE.
 DR InterPro: IPR003141; PHP.N.
 DR InterPro: IPR002309; tRNA-synt_2.
 DR Pfam: PF02231; PHP.N; 1.
 DR Pfam: PF01336; tRNA_ant1; 1.
 DR SMART: SM00481; POLIITAC; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW Complete proteome.
 SQ SEQUENCE 1115 AA; 125349 MW; E8B04E339E512FE CRC64;

Query Match 4.0%; Score 94; DB 1; Length 1115;
 Best Local Similarity 21.7%; Pred. No. 13;
 Matches 100; Conservative 62; Mismatches 155; Indels 144; Gaps 26;

OY 20 AIOQKAYPN-PVAFVDEVREND-----LGQDNELPIDVQSATQSASTDTAN- 66
 DB 437 ADQLAKLIPSRGWTLDARQOQSPOLDRLRESSLQGVYSTARKIEPLPRHASTHAGV 496
 OY 67 -----PLDEHEBELYTT--ALENKTMINCASALNODIMRLACYDPLVHGEPAY 113
 DB 497 VLSEBPLDVVPLRGHGHIYLTQYAMDH---LEDLGLKMDPGLRLN-PLISEIT-SM 551
 OY 114 IKTGRSIRLDETIWQITKGRPOVYQETTPFELMGNEKMLTKKDAK--OLEYAA--- 167
 DB 552 IEKENIKIDLS-----SISYSD--DKTF-----SLLSKGDYTGIOLESAGMRS 594
 OY 168 --KQFTPLSLFDLRNNTPLMSSRPHNPMVYLPFEMHGNRSP--NTPSHANRQFTPNE 224
 DB 595 VLKRLKPSGLE-DIVAANA---LYRP-GPMENITPLEIDRKHGRAVHYRPHEDLRSLIEDT 649
 OY 225 F-----RAPELKFOVSVYVKAEDLMGTDSLWEGYTOQSHMQ 262
 DB 650 YGVITYGQIMMISRMAGFSLGEADLLRAVSKKKKILD-----RRSHF- 696
 OY 263 IFNGKNSRPFVHDYQPEIFLTQPYSDLPMDGKRYMIGMGAHVHNSGSAKLSRSMRA 322
 DB 697 -VEGLAKKEYGV-DTANENVY-----DLIVKFNANYG-----FNRSHAVA 732
 OY 323 YLMAGMEKKNLTVMRIMGRIFKESGSGQPDNDPDLIDYGV---GDVRFYLOLENNSN 378

```

DB 733 YSNIQCO---LAVLKAHYPLFCMLTFSVIGNEKDISQVLYAKSGIRLPPSVNKS 789
QY 379 I-----SGTVRVPNSGKALQLDIVYPLGKISGFQIFQ 414
DB 790 FPETVNGSVYRAIKSV-----GVSAYKDIYK 819

RESULT 12
SRK6_BRAOL STANDARD: PRT; 849 AA.
ID SRK6_BRAOL STANDARD: PRT; 849 AA.
AC 009092;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
DE (S-RECEPTOR KINASE) (SRK).
GN SRK6.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. S656; TISSUE=Stigma;
RC MEDLINE=92020942; PubMed=1681543;
RA Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
RT "Molecular cloning of a putative receptor protein kinase gene encoded
RT at the self-incompatibility locus of Brassica oleracea."
RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
CC -1- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
CC (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
CC CYTOPLASMIC DOMAIN.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHER.
CC -1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
CC B. OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
CC -1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: M76647; AAA33000.1; ALT_TERM.
DB HSSP: P1362; IFGI.
DR InterPro: IPR001480; B_lectin.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR InterPro: IPR000858; Slocus_glycop.
DR Pfam: PR00069; pkinase; 1.
DR Pfam: PR00954; Slocus_glycop; 1.
DR SMART: SM00108; B_lectin; 1.
DR SMART: SM00473; Pan_LAP; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Transfaser: Serine/threonine-protein kinase: Signal; ATP-binding;
KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
FT SIGNAL 1 32
FT CHAIN 33 849 PUTATIVE SERINE/THREONINE KINASE

```

```

FT FT DOMAIN 33 446 RECEPTOR.
FT TRANSMEM 447 466 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 467 479 POTENTIAL.
FT NP_BIND 528 779 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 534 542 PROTEIN KINASE.
FT BINDING 556 556 ATP (BY SIMILARITY).
FT ACT_SITE 653 653 ATP (BY SIMILARITY).
FT CARBOHYD 47 47 BY SIMILARITY.
FT CARBOHYD 120 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 849;
Best Local Similarity 21.9%; Pred. No. 12;
Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;

QY 105 LVHGETPAVIRKTRSRIRLDETTIQTIGKPOVYVYQETTDPIFLMG---NEK---GMLTK 157
DB 24 LIHPALSIYINT-----LSSTESLFTISSNKTIV---SPGSIFFVGFPRTRNSRWYLGMYK 75
QY 158 K-DAKOLEYAAKOPTPLSLSF-----LDRNNTPLNSRRPNNPVYVLPFMHCK 205
DB 76 KVSDDRYVVAANDPNPLSNIGFLKISGNLVLDHSNKKVMMNTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAROPT-----PNEFRABELKFOVSKVKAEDLW 245
DB 126 -ERSPVYAEILLANGNPFMRSSNNNDASEYIMOSEPDPTDILEMKLGMYLKT----- 177
QY 246 GTSDLMFGYTGQSHWQIFNGKNSRPFVHDYQ-----PEIFLQ---PVYSDLPWDGK 296
DB 178 GLNRFL-----TSWRSSDPSSGNF---SYKLEQSLPEFLYSRENFPMHRSRPGWG- 226
QY 297 VRHIGKAVH-----HNGESAKLSRSNNRAY-----LMAGEMKNLVMF--RI 339
DB 227 IRRSGIPEDDKLSYMYNFLENNEEVAAYTFPMTNNSFSLTLTISEYFQRLTWYPSIRI 286
QY 340 WGRIFRFGSGSPDPDNDIDY-----YGYGVRFVLTQLENKSNISGTV-VYPRS 389
DB 287 WNRFW-----SSPYD-PQCDTYIMCGPYATCDV-----NTSPVCNCIOGFNPRN 329

RESULT 13
MET_MOUSE STANDARD: PRT; 1379 AA.
ID MET_MOUSE STANDARD: PRT; 1379 AA.
AC P16056; O62125;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (MET PROTO-
DE ONCOGENE TYROSINE KINASE) (C-MET) (HGF RECEPTOR) (HGF-SF RECEPTOR).
GN MET.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=88262253; PubMed=2838789;
RA Chan A.M.L., King H.W.S., Deakin E.A., Tempest P.R., Hilkens J.,
RA Kroezen V., Edwards D.R., Wills A.J., Brookes P., Cooper C.S.;
RT "Characterization of the mouse met proto-oncogene."
RL Oncogene 2:593-599(1988).
RN [2]
RP SEQUENCE OF 1199-1270 FROM N.A.
RC MEDLINE=90152381; PubMed=2482828;
RA Wilks A.F., Kurban R.R., Hovens C.M., Ralph S.J.;
RT "The application of the polymerase chain reaction to cloning members

```

RT of the protein tyrosine kinase family.";
 RL Gene 85:67-74(1989).
 CC -1- FUNCTION: RECEPTOR FOR HEPATOCYTE GROWTH FACTOR. HAS A TYROSINE-
 CC PROTEIN KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN
 CC TYROSINE PHOSPHATE
 CC -1- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN (50 KDA) AND A BETA
 CC CHAIN (145 KDA) WHICH ARE DISULFIDE LINKED.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- DISEASE: ACTIVATION OF MET AFTER REARRANGEMENT WITH THE TPR
 CC (TRANSLOCATED PROMOTER) LOCUS OF CHROMOSOME 1 PRODUCES AN
 CC ONCOGENIC PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE MET TYROSINE KINASE FAMILY OF
 CC RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00671; CA668680.1; -;
 DR EMBL: M33424; AAA40015.1; -;
 DR PIR: S01254; S01254.
 DR HSSP: P11362; 1FG1.
 DR MGD: MGI:96969; Met.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002909; IPT_TIG.
 DR InterPro: IPR003659; PSI.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF01437; Plexin_repeat; 1.
 DR Pfam: PF01403; Sema; 1.
 DR Pfam: PF01833; TIG; 3.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00429; IPT; 4.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 1379 HEPATOCYTE GROWTH FACTOR RECEPTOR.
 FT DOMAIN 25 931 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 932 934 POTENTIAL.
 FT DOMAIN 935 1379 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 55 499 SEMA.
 FT DOMAIN 1076 1343 PROTEIN KINASE.
 FT SITE 306 307 CLEAVAGE (POTENTIAL).
 FT BINDING 1082 1090 ATP (BY SIMILARITY).
 FT BINDING 1108 1108 ATP (BY SIMILARITY).
 FT ACT_SITE 1202 1202 BY SIMILARITY.
 FT MOD_RES 1233 1233 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 784 784 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 878 878 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 1199 1199 V -> I (IN REF. 2).
 FT CONFLICT 1255 1255 V -> R (IN REF. 2).
 FT CONFLICT 1261 1261 K -> T (IN REF. 2).

FT CONFLICT 1269 1270 VL -> IP (IN REF. 2).
 SQ SEQUENCE 1379 AA; 153548 MW; FC5CC87FDD8ADE8 CRC64;
 Query Match 3.9%; Score 92.5; DB 1; Length 1379;
 Best Local Similarity 21.3%; Pred. No. 24;
 Matches 80; Conservative 48; Mismatches 123; Indels 125; Gaps 22;
 QY 39 SENDGADNELPT-DVQASATOSASTDTANPDHEHRELYTTALENKTMLINCALSALNODIM 97
 DB 1014 SQNGACRQVQPLDLSPLILNSGSDISSPL-----LQN-TYHIDLSALNELV 1061
 QY 98 RLACTDVLHGETPAVITKTSIRLDETITWOTIKGPOVYQETDPIFLMGNEKGMTLK 157
 DB 1062 QAVQH-VVIGPSSSLI-----VHEVEVIG-RGHFCVYV-----GTLID 1097
 QY 158 KDAKLEVAARQFPLSLSPDLDRNPTPLMS-----SRP-----NPVY 197
 DB 1098 NDGKRIHCAVNSLNRIIT---DIEVSOFLTEGITIKKDSHPNVLSLGLCLRSQSPLYV 1154
 QY 198 LPIFMHGKPNNSPNTPSHEARQFPNEFRABELKQVSVKKAADLWGTDSDLMFGYTO 257
 DB 1155 LPYMKHG-----DLRNFIRETHNPYVK-----DLIG-----FGL- 1184
 QY 258 OSHWOIFNGKN---SRPRVADYQPEIFLTQPVYSDELPMDCGVNRIGM-----AV 305
 DB 1185 ---QVARGMKYLAASKKFEVHRD---LAARCMIDKEFTVAVDFGLARMDYKEYYSV 1235
 QY 306 HHSNGESAKLSRSMRAVLMGMEMKNTLVMPRING-----RIPEKSGSQPDNDP-DI 358
 DB 1236 HNKTS-AKLEPVKM---MALESLOQKFTTSDVMSFVGLMELMTRARPRPDVNTDI 1290
 QY 359 LDYGYGVDFRPLYOLE 374
 DB 1291 TTYLQG---RRLLQPE 1304
 RESULT 14
 YL70_ARCFU STANDARD; PRT; 324 AA.
 ID YL70_ARCFU
 AC 028112;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN AF2170.
 GN AF2170.
 OS Archaeoglobus fulgidus.
 CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 CC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Richardson D.L., Kierlavage A.R., Graham D.B., Kyriides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
 RA Cotton M.D., Spriggs T., Artlach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- SIMILARITY: CONTAINS 2 KELCH REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:27 ; Search time 77.9 Seconds

(without alignments)
829,941 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSLSCFALLA.....YNHEATSPFGVGLMDNMGL 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2 09K007	09K007 neisseria m
2	815	34.5	409	2 09J021	09J021 neisseria m
3	373.5	15.8	306	2 09C122	09C122 pasteurella
4	366	15.3	329	2 09PM08	09PM08 campylobact
5	360.5	15.3	292	2 09S1D7	09S1D7 yersinia ps
6	360	15.3	289	2 09L6N9	09L6N9 salmonella
7	356.5	15.1	292	2 09Z4N8	09Z4N8 enterobacte
8	342.5	14.5	297	2 032349	032349 campylobact
9	246.5	10.4	355	2 025241	025241 helicobacte
10	240.5	10.2	355	2 09ZLX5	09ZLX5 helicobacte
11	123	5.2	278	2 09XB53	09XB53 erwinia car
12	114	4.8	1686	4 000443	000443 homo sapien
13	110.5	4.7	602	11 063485	063485 rattus norv
14	109.5	4.6	821	2 059241	059241 bacillus sp
15	108	4.6	824	2 09F216	09F216 bacillus sp
16	105.5	4.5	901	5 018749	018749 caenorhabdi
17	104	4.4	783	2 045554	045554 bacillus sp
18	104	4.4	798	4 09UN32	09UN32 homo sapien
19	104	4.4	798	4 09UBK2	09UBK2 homo sapien

20	103.5	4.4	3247	12 065553	065553 bovine herp
21	103	4.4	435	2 09X9C0	09X9C0 streptococc
22	101.5	4.3	403	5 P91736	P91736 hydra magni
23	101.5	4.3	1046	2 084941	084941 streptococc
24	101	4.3	4307	5 019319	019319 caenorhabdi
25	100.5	4.3	1509	11 061194	061194 mus muscubu
26	100	4.2	719	2 057114	057114 streptococc
27	100	4.2	719	2 09REM8	09REM8 streptococc
28	99.5	4.2	739	5 09V9B6	09V9B6 streptococc
29	99.5	4.2	1658	11 061182	061182 mus muscubu
30	99	4.2	719	2 054948	054948 streptococc
31	99	4.2	719	2 054946	054946 streptococc
32	99	4.2	719	2 09WVW0	09WVW0 streptococc
33	99	4.2	719	2 09WVW1	09WVW1 streptococc
34	99	4.2	719	2 09RET6	09RET6 streptococc
35	99	4.2	719	2 09RET4	09RET4 streptococc
36	99	4.2	727	2 09F2G4	09F2G4 streptococc
37	99	4.2	1372	10 09FL92	09FL92 arabidopsis
38	99	4.2	5005	2 09P2P5	09P2P5 ureaplasma
39	98.5	4.2	467	4 09UJMS	09UJMS homo sapien
40	98.5	4.2	467	4 09BYT2	09BYT2 homo sapien
41	98.5	4.2	576	2 09L115	09L115 streptomyc
42	98.5	4.2	660	2 09KGR8	09KGR8 bacillus ha
43	98.5	4.2	788	10 09SCV5	09SCV5 arabidopsis
44	98	4.2	719	2 054947	054947 streptococc
45	98	4.2	719	2 054949	054949 streptococc

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	382 AA.
09K007	09K007			
ID	09K007			
AC	09K007			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC58 / SEROGROUP B;			
RX	MEDLINE=20175755; PubMed=10710307;			
RA	Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,			
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,			
RA	Nelson W.C., Gwyn M.L., DeBoy R., Peterson J.D., Hickey E.K.,			
RA	Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,			
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Ciltone H., Clark E.B.,			
RA	Cotton M.D., Utechtack T.R., Khouri H., Qin H., Yamahavean J.,			
RA	Gill J., Scarlato V., Maignani V., Piazza M., Grandi G., Sun L.,			
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;			
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain			
RT	MC58."			
RL	Science 287:1809-1815(2000).			
DR	EMBL: AE002403; AAF40901.1; -.			
DR	TIGR: NMB0464; -.			
DR	InterPro: IPR003187; PLA1.			
DR	Pfam: PF02253; PLA1; 1.			
KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;			
Query Match	34.5%; Score 815; DB 2; Length 382;			
Best Local Similarity	44.7%; Pred. No. 1,66-62;			
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;				
QY	87 INCSALNDINRLACVDTLVGERTPAVI-----KTKRSIRLDETIMQTI-KGKPVVYQOE 140			
DB	33 LQCAALFTDNTVRLACVDRIFAQALPSSAGGEGESKAVLNTETVRSLSLDGEAVIYVEK 92			


```

DE PHOSPHOLIPASE A (EC 3.1.1.32).
GN PLDA OR C1351.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCITC 1168.
RC MEDLINE=20150912; PubMed=10689204;
RA Parhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jags K., Kariyev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Owall M.A., Raftery M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL: AL13078; CAB7378.1; -.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
KW Complete proteome.
SQ SEQUENCE 329 AA; 38880 MW; A32731F2B751AC44 CRC64;

Query Match 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 1.2e-23;
Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

OY 130 IKGKPOVVOETTDPIFLMGNEKMKLTKKDAQOLEVAAQKQFPLSLSPDNRNT--PLW 187
DB 43 LKSSSVLIQEQONSSQATQTONSITIKKEKO-----DFSLALANTIGENESRPLG 96
OY 188 SSRPHNPVLPFLFMHGKPNRSPNPSHEAROPTNEFRAPELKFOVSVKAAEDLMGT 247
DB 97 IS-SYKMYFLP-FAYSFNSLGVNNKSEA-----KFLQSVYKRLFENLGL 141
OY 248 DSDLMFGYQOOSIMQIFNGKNSRPFVHDYQPEIFLTQPVY-SDLPMWDGVNRIGKAVH 306
DB 142 DEKYLAIVQTSWQIVY--EHSSPFRETNYQPFIDLPYLKDFEFNNLR--VGILH 196
OY 307 HSNGESAK--LSRNNRAYLAMGEMKNTLVMPRIKRIKESGSGSDPDNDPILDYGY 364
DB 197 ESNKGDEMLQSSMNRIVSTAILNKFLFVRLMYRT--PENKKDDNPAILHYMGN 253
OY 365 GGVRLYQLKNSNISGVARYNPR--SGKALQLDYVPL-GKISGYFOIFOGYQSOLI 421
DB 254 FQVNLAY-LGDDYFINMLRLNNLKFNHNGAIGVDIGDIFNGGIWYLQYFNGYGSOLI 312
OY 422 DYNHEATSFVGLMLN 437
DB 313 DYNKHLQRLSTGFLIS 328

RESULT 5
OQSID7 PRELIMINARY; PRT: 292 AA.
AC OQSID7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Versinia pseudotuberculosis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Versinia.
OX NCBI_TaxID=633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PIIT PIBI;
RC Kariyev A.V., Winzeler E.A., Williams K.J., Oyston P.C.,
RA Tildall R.W., Wren B.W.;

```

```

RT "Biochip-based Signature-Tagged Mutagenesis: Identification and
RT characterisation of V. pseudotuberculosis gene plda essential for
RT virulence in mice."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ245393; CAB51586.1; -.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
SQ SEQUENCE 292 AA; 33758 MW; 8E712D908ACB5A5 CRC64;

Query Match 15.3%; Score 360.5; DB 2; Length 292;
Best Local Similarity 39.2%; Pred. No. 3e-23;
Matches 83; Conservative 34; Mismatches 86; Indels 9; Gaps 5;

OY 229 ELKFQVSVKKAEDLMGTDSDLMFGYQOOSIMQIFNGKNSRPFVHDYQPEIFLTQPV- 287
DB 86 EVKFOLSIAPPIWRGIGDSSLGASVYQSRSMQWQNSSESSPFRETNYEPQLFLAWSTD 145
OY 288 YSDLPMDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLAMGEMKNTLVMPRIKRIKES 347
DB 146 YELAGW--TFREVEFEGFNHOSNGKADPTSRSMNRYRVAQKGNLEIDLPWRIPESD 203
OY 348 SSGSDPDNDPILDYGYGADVFLYQL-ENKSNISGVARYNPRSGKALQLDYVPLGKI 406
DB 204 S---KDDNPITKMGYRLKVGATLGDVSFSLDG--RYWMNMGYGAEMGMSYPIRKHV 258
OY 407 SGYFOIFOGYQSOLIDYNHEATSFVGLMLND 438
DB 259 RFTYQVFSYGESMIDYFNQTRVGVIMLND 290

RESULT 6
O9L6N9 PRELIMINARY; PRT: 289 AA.
AC O9L6N9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE PLDA PROTEIN.
GN PLDA.
OS Salmonella typhimurium LT2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=99287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SGSC1412;
RA WashU;
RT "The Salmonella typhimurium Genome Sequencing Project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SGSC1412;
RA Waterston R.;
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF233324; AAF33435.1; -.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
SQ SEQUENCE 289 AA; 32967 MW; DA97F5E1651C49C6 CRC64;

Query Match 15.3%; Score 360; DB 2; Length 289;
Best Local Similarity 39.7%; Pred. No. 3.3e-23;
Matches 85; Conservative 29; Mismatches 90; Indels 10; Gaps 4;

OY 226 RAELEKFOVSVKKAEDLMGTDSDLMFGYQOOSIMQIFNGKNSRPFVHDYQPEIFLTQ 285
DB 83 KDEVYKFOLSIAPPIWRGIGDSSLGASVYQSRSMQWQNSSESSPFRETNYEPQLFLG 142
OY 286 PV-YSDLPMDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLAMGEMKNTLVMPRIKRI 344
DB 143 ATDYRFAGW--TLRDVEMGYNHDSNGRSDPTSRSMNRLYRLMAENGNNMLVEYVPMYVI- 189

```

QY	345	KEGSSGQPDNDGDIIDYGYGSDVRFRLYOLEKNSNSIGTRYRNPBSRGKALQIDYVYPLCK	404
Db	200	-----GSTDNDNDITRYKMYQKLTGKGYHL-GGAVALSARQOYNMNTGYGGAEGLSYPIYK	253
QY	405	GISGYFOIFOGYGOSLIDYNHEATSEFGVGLMND	438
Db	254	HYRLTYQVYSGYGESLIDYENFQNTRVGCVMLND	287
RESULT	7		
Q924N8	0924N8	PRELIMINARY:	PRT: 292 AA.
AC	0924N8:		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (Ec 3.1.1.32).		
GN	PLD.		
OS	Enterobacter agglomerans.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OX	Pantoea.		
RN	NCBI_TaxID=549;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=94131966; PubMed=8300539;		
RA	Brok R.G., Brinkman E., van Bostel R., Bekkers A.C., Verheij H.M.,		
RA	Tommassen J.;		
RT	"Molecular characterization of enterobacterial plid genes encoding		
RT	outer membrane phospholipase A.;"		
RL	J. Bacteriol. 176:861-870(1994).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99120145; PubMed=9921577;		
RA	Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tommassen J.;		
RT	"Sequence comparison of outer membrane phospholipases A: implications		
RT	for structure and for the catalytic mechanism.;"		
RL	Res. Microbiol. 149:703-710(1998).		
DR	EMBL: AF034414; AAD03498.1; "		
DR	InterPro: IPR003187; PLAI.		
DR	Pfam: PF02253; PLAI.1.		
DR	PRINTS: PRO1486; PHPLIPASEAL.		
KW	Signal; Hydrolase.		
FT	SIGNAL	1	20
FT	CHAIN	21	292
FT	SEQUENCE	292 AA;	33719 MW; B75516D093B2BEA CRC64;
FT	OUTER MEMBRANE PHOSPHOLIPASE A.		
QY	Query Match	15.1%;	Score 356.5; DB 2; Length 292;
QY	Best Local Similarity	40.1%;	Pred. No. 6.8e-23;
QY	Matches	85; Conservative	31; Mismatches 87; Indels 9; Gaps 5;
QY	229	ELKPVASVKVKAEDLMDGTDSDLMTCGYQOSHWQIFNGKNSPRFVHYDQPEIFLTQPV-	287
Db	86	EKKFQLSLGFPIMRGIAADNSLGLASYQSRWQASNSDESSPFEETNPEQIFLAMATD	145
QY	288	YSDLPMDCKVRMIGAGVAHNSGESAKLSRSWNRAVYLAAGMEKMKILTVAPRTMGRTFKEG	347
Db	146	YELAAW--TFREVEEGYNHQNSGKADPISRSMDRYITRLMQRGNLEIDIKRWYRIPESD	203
QY	348	SGSQPDNDPDIIDYGYGDVRFRLYOL-ENKSNISGTVRYRNPBSRGKALQIDYVYPLGKI	406
Db	204	S---KDDNPDIIDIKYGYRLKGYALGSEVSFIDG--RYNMNTGYGGAEMGMSYITKHV	258
QY	407	SGYFOIFOGYGOSLIDYNHEATSEFGVGLMND	438
Db	259	RFYTVFSGYGESMTDYNFRQTRVGVGLMND	290
RESULT	8		
Q32349	032349	PRELIMINARY:	PRT: 297 AA.
ID	032349:		
AC	032349:		

Query Match	14.5%	Score 342.5	DB 2	Length 297
Best Local Similarity	32.4%	Pred. No. 1.1e-21		
Matches 97	Conservative 37	Mismatches 108	Indels 57	Gaps
01-JAN-1998 (TReMBLrel. 05, Created)				
01-JAN-1998 (TReMBLrel. 05, Last sequence update)				
01-JUN-2001 (TReMBLrel. 17, Last annotation update)				
PHOSPHOLIPASE A.				
PLDA.				
Campylobacter coli.				
Bacteria: Proteobacteria; epsilon subdivision; Campylobacter group.				
Campylobacter.				
NCBI_TaxID=195;				
SEQUENCE FROM N.A.				
STRAIN=UA585;				
MEDLINE=97230284; PubMed=9119448;				
Grant K.A., Belandria I., Dekker N., Richardson P.T., Park S.F.,				
"Molecular characterization of plda, the structural gene for a				
phospholipase A from Campylobacter coli, and its contribution to cell-1-				
associated hemolysis";				
Infect. Immun. 65:1172-1180(1997).				
EMBL: Y11031; CAA71915.1; -				
InterPro: IPR003187; PLAI.				
Pfam: PF02253; PLAI.1.				
PRINTS: PRO1486; PPHPIIPASEA1.				
SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;				

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.; Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 388:539-547(1997).
 DR EMBL: AE000564; AAD07564.1; -
 DR TIGR: HP0499; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 3; 2e-13;
 Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

OY 157 KDAKOLEYAAKQFTPLSFDDLRRNTPLMSSRPHNPVVLPIFMHGKPNRSPNPSHE 216
 DB 69 KKLNMMDVIGTYFLPRYHSF-----TFIFQYHNPINP----- 102
 OY 217 AROFTPNEFRAPLKFQVSVKAAEDLWGTSDLMFGYQOSHWOIFNGKNSRPFVVD 276
 DB 103 ---YORNEF-----KFOISFRVPEVFRHILWTGTLAYLTQDWFQYNDPOSAPMRM 154
 OY 277 YOREIFLTQPVYSDLPMDGV---RMIGMAVHNSG--ESAKLSRSNRAYLAGMEKRN 332
 DB 155 FMEPLIYVYPI-NKPRFGKIGNFSIWMQHISNGVGAOCYOPNPK---EGNPDNQ 209
 OY 333 LTVMPRI-----MGRIFEKSGSQP-----DDNPD 357
 DB 210 FPGQPIVADYNGQKDYRWGCGRSVSAGQRPVRLVWEKGLKIMVAYWYVYDQSNPN 269
 OY 358 ILDYGYGVDFLY-----OLENKSNIIGTVRNPBSGKALQLDYVYPLGKISGYF 410
 DB 270 LIDYMGYGNAKIDYRGRHNFELQLDYDFQYWRD--RWHGAFRLGYTYRINPFV 327
 OY 411 QIFQGYGOSLIDYNHATSFGVGLMLN 437
 DB 328 QWENGYGDLGYEYDFVSNRIGVIRLN 354

RESULT 10
 O9ZLX5 PRELIMINARY; PRT; 355 AA.
 AC O9ZLX5;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PUTATIVE PHOSPHOLIPASE A1.
 GN PLDA OR JHP0451.
 OS *Helicobacter pylori* j99 (Campylobacter *pylori* j99).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OC *Helicobacter*.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Molt D.T., King B.L., Brown E.D., Dolg P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RL Nature 397:176-180(1999).
 DR EMBL: AE001479; AAD06029.1; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Complete proteome.
 SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 1e-12;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

OY 157 KDAKOLEYAAKQFTPLSFDDLRRNTPLMSSRPHNPVVLPIFMHGKPNRSPNPSHE 216
 DB 69 KKLNMMDVIGTYFLPRYHSF-----TFIFQYHNPINP----- 102
 OY 217 AROFTPNEFRAPLKFQVSVKAAEDLWGTSDLMFGYQOSHWOIFNGKNSRPFVVD 276
 DB 103 ---YORNEF-----KFOISFRVPEVFRHILWTGTLAYLTQDWFQYNDPOSAPMRM 154
 OY 277 YOREIFLTQPVYSDLPMDGV---RMIGMAVHNSG--ESAKLSRSNR----- 321
 DB 155 FMEPLIYVYPI-NKPRFGKIGNFSIWMQHISNGVGAOCYOPNPKGNPDNPPQ 213
 OY 322 -----ATLMAG-----MEWK--NLTVMPRIGRIFKESGSGQPPD- 354
 DB 214 PVIYKDYNGQKDYRWGCGRSVSAGNALCFVLMWEKGLKIMVAYWYV-----PYDQ 265
 OY 355 -NPDIIDYGYGVDFLY-----OLENKSNIIGTVRNPBSGKALQLDYVYPLGKGT 406
 DB 266 SNPQLIDYMGYGNAKIDYRGRHNFELQLDYDFQYWRD--RWHGAFRLGYTYRINPFV 323
 OY 407 SGYFOIFQGYGOSLIDYNHATSFGVGLMLN 437
 DB 324 GIYAQWENGYGDLGYEYDFVSNRIGVIRLN 354

RESULT 11
 O9XB53 PRELIMINARY; PRT; 278 AA.
 AC O9XB53;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE KDOT.
 GN KDOT.
 OS *Erwinia carotovora*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_TaxID=554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048; GS101.
 RA McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,
 RA Stewart G.S., Bycroft B.W., Salmood G.P.;
 RT "Analysis of the carbenem gene cluster of *Erwinia carotovora*:
 RT definition of the antibiotic biosynthetic genes and evidence for a
 RT novel beta-lactam resistance mechanism.";
 RL Mol. Microbiol. 26:545-556(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048; GS101.
 RX MEDLINE=98276484; PubMed=9614345;
 RA McGowan S.J., Bycroft B.W., Salmood G.P.;
 RT "Bacterial production of carbapenems and clavams: evolution of beta-
 RT lactam antibiotic pathways.";
 RN Trends Microbiol. 6:203-208(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048; GS101;
 RA McGowan S.J.;
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U17224; AAD38237.1; -
 SQ SEQUENCE 278 AA; 31341 MW; 63769FAF3550E1B7 CRC64;

Query Match 5.2%; Score 123; DB 2; Length 278;
 Best Local Similarity 21.6%; Pred. No. 0.012;
 Matches 74; Conservative 52; Mismatches 128; Indels 88; Gaps 16;


```

0Y 29 NPAVAVDEVRASNDQGOENELPIDVQASOTDASTDPAJNDEHBEPEYTTALEKMTLIN 88
Db 87 NOVAVUTKKNLELEAODRNIGT-----QOFTRAKEELEAEKRDILRT---NERLSOE 133
0Y 89 CSALNODIMRLACVDTLVHGETPAVITKTKRSIRIDETIMO---TIKRPVOYVOE----- 140
Db 138 VEYLEDYVKRL--NEKTESNT---TKGELQTKIDELQASDVYKREKRLGCKELLHN 192
0Y 141 -----TTDPIFLMGNEG-----MLTKKDAQOLEYAAKQTPPLSLSPDLD 180
Db 193 ONSWLNTELKTKTBEALLGREGKEWELTELKCTLENKEEDALISHSESAPSALS--SSP 251
0Y 181 RNNTPLMRSSRPHNPYUVALPIFMHGKPNRSPTPSHEARQTPNEFRADELFOYSVYKA 240
Db 252 NNLSPGTGSOQKTP-----VPRAOREAPSSGOEKKIKRPROGROSSYWELEASEVM 300
0Y 241 AEDLMGTSDLMFGYTOOSH-----QIFGNKSRPREVHDYPER----- 281
Db 305 LSTRIGSGS---FGVUYGKGNHGDVAIVILKVUPPTPEOLQAFNEVAYLEKTRHVVLL 361
0Y 282 -----FLPO-----PVYSDL--PMDGKVMI-----GMGAVHNSGESA 313
Db 362 FMGYTKTKNLALVYOMCGSSLYKHLHVOETKRFOMFOULIDAROTAOGMUYLHAKNIIR 422
0Y 314 KLSRSMNBAVYLMAGMEK-----NLTVMPRIWGRIFEGESGSO----- 351
Db 422 DMKS--NNIFLHEGLTVIGDFGLATVKSRY-----SGSQVEOPTGSVLYMAPEYIR 472
0Y 352 -PDDNPDLL--DYVCYGDVRFYLXOLEKNSNSGIVRVNPRSGKALODUYVPLGK 405
Db 473 MODNPNPFSQSDVSYGLV--LYEL-----MTGELPFSHINNROI-----IFWNGK 518

```

```

RESULT      14
OS9241
ID      OS9241      PRELIMINARY;      PRT;      821 AA.
AC      OS9241;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      ENDO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULOSE) (ENDOGALACTANASE)
DE      (CARBOXYMETHYL CELLULOSE).
OS      Bacillus sp.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
OC      Bacillus/Staphylococcus group; Bacillus.
OX      NCBI_TaxID=1409;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-KSM-64;
RA      Sumitomo N., Ozaki K., Ito S.;
RL      Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC      -1- CATALYTIC ACTIVITY: ENDOMYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC      LINKAGES IN CELLULOSE.
CC      EMBL; M84963; AAAT3189.1; -.
DR      HSSP; O85465; 1A3H.
DR      InterPro; IPR001547; Glyco_hydro_F5.
DR      Pfam; PF00150; cellulase.1
DR      PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWNM.1.
KW      Hydrolyase. Glycosidase.
SQ      SEQUENCE      821 AA; 90910 MW; 73DA38FEF0B40B5C CRC64;

```

Query Match	4.68;	Score 109.5;	DB 2;	Length 821;
Best Local Similarity	19.88;	Pred. No. 0.81;		
Matches 95;	Conservative 54;	Mismatches 163;	Indels 169;	Gaps 25

```

0y 9 TLSIISCAIILAIQAKAVPMPVAFVDEVEREND---LGOENELPIVQSTQASIDTA 65
    | : : : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 TKQISLSLILVL-TLSLFPALAAEGNTRDNRKHLGLGNVYKRPSEAGALQIQEYDQG 64
    | : : : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
0y 66 NPL-DEHEPELY-----TTALENTMLINCALS-----NQDMLRACY-DLIVHGEP 111
    | : : : | | : | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db      65  MTLVDQHEKXIQGLQGMSTHGLQWPELINDNAKYALANDMESNMIRLAMYVGENGASNP 124
QY      112  AVIKTKRSIRLDETI-----WQT-----IKGPOVY 138
           : : : : :
Db      125  ELIKSRVIKGIDLAIENDMYIVDMVHAFDPDRDPYAGAEDEFPRDIALYNNPHIY 184
QY      139  QETIDP-----IFLKGNEKGMLTKKDQOLEYAAKOFPLSLSPDLN-----NPLW 187
           : : : : :
Db      185  ELANEPSSNNNGAGIRPNEEGMAVYK-----EYADPIVELLRSGNADNIIIVGSPNW 239
QY      188  SSRP-----HNPIYVLPiF--MHGKPNRS--FNTSHEAKOTTPMEFARPELKFO 233
           : : : : :
Db      240  SORPDLADNPIIDHHTWTYVHAFYTGSHASTESYPEPTEPSEKGNVMSNTRYA-----LE 295
QY      234  VSVKKAADLMGT-----DSDDLFGYTOQ-----SHWQIFNCKNS-----RP 271
           : : : : :
Db      296  NGVAVFATE--WGTSGQANGDGPPFDEADWIEFLNENNISWAMSLTN--KNEVSGAFTP 352
QY      272  FRVH-----DROPELFLQPYVSDLPW-----D 294
           : : : : :
Db      353  FELGKSNAATSLDPCGPDQVWPEELSLSGEYVRARIKGVNERP--IDRTKYKVLWEND 409
QY      295  GKVRMIGMGAVHSHNGES-----AKLSRSMNRAYLTLAGMEKMLTVMPIRW 341
           : : : : :
Db      410  GTKQGEFV-----NGDSPEVEDVYIENEGALKLTLGSLDASNDVSEGVYMANARLSADGWG 463
QY      342  R 342
Db      464  K 464

```

RESULT	15		
ID	Q9F216	PRELIMINARY;	PRT; 824 AA.
AC	Q9F216;		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	CELLULOSE.		
OS	Bacillus sp.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
OX	NCBI_TaxID=1409;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=KSM-5237;		
RA	MEDLINE=21036886; PubMed=11193393;		
RA	Hakemada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,		
RA	Kobayashi T., Ito S.;		
RT	"Purified amino acid sequence and possible catalytic residues of a		
RT	thermostable, alkaline cellulase from an alkaliphilic Bacillus		
RT	strain.";		
RL	Biosci. Biotechnol. Biochem. 64:2281-2289(2000).		
DR	EMBL; AB018420; BAB19360.1; -		
DR	InterPro: IPR001547; Glyco_hydro_F5.		
DR	Pfam: PF00150; cellulase.1.		
DR	PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN.1.		
SO	SEQUENCE 824 AA: 91564 MW; 65FA940FE1DD729E9 CRC64;		

Query Match	4.6%;	Score 108;	DB 2;	Length 824;
Best Local Similarity	19.8%;	Pred. No. 1.1;		
Matches	95;	Conservative	60;	Mismatches 159;
				Indels 166;
				Gaps 28;

[illegible]

```
Db 122 ATNPBLIKQRIYDGIETLAIENDMYIVDMHVHAPGDDPRDPVYAGAKDFPREIAALYPNNP 181
QY 135 QVVOETTP-----IFLMGNEKMLTKKDAKOLEYAAKQFTPLSFDLDRN-----N 183
Db 182 HIIFELANEPSSNNNGAGIPNNEGKAVK-----ETADPIVEMLRKSGNADNIIYVG 236
QY 184 TPLMSSRP-----HNPMYVLPIFMHGKPNRSPNTPSHANQFTPNEFRAPELK-- 231
Db 237 SPNMSQRPDLADNPIDDHHTWYVHFYTGSHAASTESYPSE-----TPNSERGNVMSNT 291
QY 232 ---FQVSVKVKAEDLMGT-----DSDLMPGYTQO-----SHWOIFNGKNS-- 269
Db 292 RYALENGVAVFATE--WGTSQASGDGCPYFDEADVWIEFLNENNISWANWSLJN-KNEYS 348
QY 270 ---RPFRV-----HDYOP-EIIFL-----QPV-----YSDLPW-- 293
Db 349 GAFTEFELGKSNATNLDGPDHVMAPPEELSLSGEYVRARIKGVNTEPIDRTKYTKVLMDF 408
QY 294 -DGKVRMIGMCAVHHSGESAKLSRSMNRAVIMAGME-----WKNLTVMPRIWGR 342
Db 409 NDGTRQ--GFGVNSDSPNKEIAYDNENNTLKVSGLDVSNVSDGNFMANARLSANGWGK 466
```

Search completed: November 30, 2001, 14:26:28
Job time: 565 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:25 ; Search time 77.9 Seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 2360

Sequence: 1 MKVSLSTLTSLTLPCEFLA.....YNHEATSPGVGLMDWML 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: SP TREMBL.17:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_protent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2 09K0U7	09K0U7 neisseria m
2	815	34.5	409	2 09J721	09J721 neisseria m
3	373.5	15.8	306	2 09C422	09C422 pasteurella
4	366	15.5	329	2 09PM08	09PM08 campylobact
5	360.5	15.3	292	2 09S1D7	09S1D7 yersinia ps
6	360	15.3	289	2 09L6N9	09L6N9 salmonella
7	356.5	15.1	292	2 09Z4N8	09Z4N8 enterobacte
8	342.5	14.5	297	2 032349	032349 campylobact
9	246.5	10.4	355	2 025241	025241 helicobacte
10	240.5	10.2	355	2 09ZLX5	09ZLX5 helicobacte
11	121	5.1	278	2 09XB53	09XB53 erwinia car
12	109.5	4.6	824	2 09F216	09F216 bacillus sp
13	109	4.5	821	2 059241	059241 bacillus sp
14	106.5	4.5	602	11 063485	063485 rattus norv
15	106	4.5	798	4 09UN32	09UN32 homo sapien
16	106	4.5	798	4 09UBK2	09UBK2 homo sapien
17	102.5	4.3	783	4 045554	045554 bacillus sp
18	102.5	4.3	1686	4 000443	000443 homo sapien
19	99.5	4.2	1509	11 061194	061194 mus musculu

20	98.5	4.2	467	4 09UMW5	09UMW5 homo sapien
21	98.5	4.2	467	4 09BRV2	09BRV2 homo sapien
22	98.5	4.2	576	2 09L115	09L115 streptomyce
23	98.5	4.2	788	10 09SCV5	09SCV5 arabidopsis
24	98.5	4.2	1658	11 061182	061182 mus musculu
25	98	4.2	435	2 09X9C0	09X9C0 streptococc
26	98	4.2	871	4 09Y5C3	09Y5C3 homo sapien
27	98	4.2	938	4 09Y5F7	09Y5F7 homo sapien
28	97.5	4.1	797	2 09RDM6	09RDM6 lactobacill
29	97	4.1	761	12 09WT88	09WT88 lt virus. o
30	96	4.1	422	2 09RC26	09RC26 streptomyce
31	95.5	4.0	389	2 P74515	P74515 synechocyst
32	95	4.0	523	4 09H856	09H856 homo sapien
33	95	4.0	739	5 09V9E6	09V9E6 drosophila
34	94.5	4.0	476	5 09B160	09B160 caenorhabdi
35	94.5	4.0	530	5 045879	045879 caenorhabdi
36	94.5	4.0	683	2 09A6U7	09A6U7 caulobacter
37	94.5	4.0	749	2 059154	059154 anaerocellu
38	94.5	4.0	772	2 0923V2	0923V2 pseudomonas
39	94.5	4.0	1046	2 084941	084941 streptococc
40	94.5	4.0	1641	2 09PDX7	09PDX7 xylella fas
41	94	4.0	454	11 09EG54	09EG54 mus musculu
42	94	4.0	455	4 012875	012875 homo sapien
43	94	4.0	765	2 054183	054183 streptomyce
44	93.5	4.0	564	3 094727	094727 schistosach
45	93.5	4.0	664	3 09C122	09C122 piromyces s

ALIGNMENTS

RESULT	1	ALIGNMENTS
09K0U7	PRELIMINARY:	PRT: 382 AA.
ID 09K0U7		
AC 09K0U7		
DT 01-OCT-2000 (TREMBLrel. 15, Created)		
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE PHOSPHOLIPASE A1, PUTATIVE.		
GN NMB0464.		
OS Neisseria meningitidis (serogroup B).		
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX NCBI_Taxid:491;		
RN [1]		
RP SEQUENCE FROM N.A.		
RC STRAIN=MC58 / SEROGROUP B;		
RX MEDLINE=20175755; PubMed=10710307;		
RA Tetteijn H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,		
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,		
RA Nelson W.C., Gwin M.L., Deboy R., Peterson J.D., Hickey E.K.,		
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,		
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B.,		
RA Cotton M.D., Utecherack T.R., Khouri H., Qin H., Yamahavean J.,		
RA Gill J., Scarlato V., Maignani V., Pizsa M., Grandi G., Sun L.,		
RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;		
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain		
RT MC58.";		
RT Science 287:1809-1815(2000).		
RL EMBL: AE002403; AAF40901.1; -.		
DR TIGR: NMB0464; -.		
DR InterPro: IPR003187; Pfam: PF02253; Pfam: 1.		
KW Complete proteome.		
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;		
Query Match	34.5%; Score 815; DB 2; Length 382;	
Best Local Similarity	44.7%; Pred. No. 6.5e-63;	
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;		
QY 87 INCALNODIMRLACVDTLVGERTPAVL-----KTKRSIRLDETITWQTI-KGKQVYVQOE 140		
DB 33 LQCAALTDNVTIRLACVDRIFAQLPSSAGCGGSGSKAVLNTETVRSLSLDGEAVIVVEK 92		

QY	141	TTDPFLWLNENKGMILTKDKADOLEAAQOFPLTSEFDLSDRNN-TPIMSSRPHNPXYLP	199
Db	93	GGDAL-----PADSAGETADITTYPLSLMTYDLDKDKDKGLGCLREHNPMYLP	139
QY	200	IFMHCKPNRSPTPSHEAR-OETPNEFRAPELKFOVSVYKAAEDLMGSDSLMEGYTOO	258
Db	140	LMYNNSPNVAPGSPRGATVDEKFGQOKRAETKLOYFSFKSLIAEDLFKTRADLMFGYTOR	199
QY	259	SHWQLEFN-GKNSRPPRVINDYOPEITFLTOPYVSDLEPMDCKVVRMIGMAVHNSGESAKLSR	317
Db	200	SDWQLYNNGRSAPERNMTDYKKEIFELTOPVKADLPFGGRLMIGAFVHQSGQSPESR	259
QY	318	SMNRAYLWAGMEWMKMLTYWPRIMGRITFEKSGSQPDNDNDIIDDYGGYGVRELYOLENKS	377
Db	260	SMNRITVMAAGMGMKGLTYIAPRYVVAFAFD-SSDK-NDNPDIADYMGCGYVKLOTRYLNDQ	317
QY	378	NISGTVRYNPRSCKGALOLDVYVPLGKGISGYFGIOFGYGSGLIDYHNBEATSEFGVGLMLN	437
Db	318	NVYSVLRALNPKTYGATIEAAYTFPIKGLKGVVRGFGHGESLIDYHNKONGIGLGMEN	377
QY	438	DMWGL 442	
Db	378	DLDGT 382	
RESULT	2		
ID	Q9J721	PRELIMINARY:	PRT: 409 AA.
AC	Q9J721:		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PURATIVE PHOPHOLIPASE.		
GN	NMA2021.		
OS	Neisseria meningitidis (serogroup A)		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=65699;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;		
EX	MEDLINE=20222556; PubMed=10761919;		
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,		
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,		
RA	Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,		
RA	Jagels K., Leather S., Mounle S., Muggall K., Quail M.A.,		
RA	Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,		
RA	Whitehead S., Spratt B.G., Barrrell B.G.;		
RT	"Complete DNA sequence of a serogroup A strain of Neisseria		
RT	meningitidis Z2491."		
RL	Nature 404:502-506(2000).		
DR	EMBL: AL162757; CAB85240.1; -		
DR	InterPro: IPR003187; Pfam: PF02253; Pfam: 1.		
DR	Pfam: PF02253; Pfam: 1.		
KW	Complete proteome.		
QO	SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;		

Query Match	34.5%	Score 815;	DB 2;	Length 409;	
Best Local Similarity	44.7%	Pred. No. 7,1e-63;			
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps	8

QY	87	INCSALNODIMKLACDYTLVHGEPFAVI-----KRRSRIRLEDTWQRI-KGKPOVVOYE	140
	:		:
Db	60	LQCALNTDQNVTRFLACYDRIIPAQNLTPSSAGQEGQESKAVNLNLETYRSSLDKGEAVIYVEK	119
	:		:
QY	141	TTDPFLFMGNEKGMLTKKDAKOLEFAAKOFTPLSLFSDIDRRNN-TPLMSSRPHNPATYLP	199
	:		:
Db	120	GGDAL-----PADSAGETADIVPLSLMYDLDKNDKGLLREHNPATYLP	166
	:		:
QY	200	IFMHGKPNRSPPTPSHEAR-OFTPNRPEPAELKFOVSVYVKAAPEDLWGDSDSLMFGYTQO	258
	:		:
Db	167	LMYNNSPNTAPROSPFRGTTGVDEKFGQQRKAETKLVQSEKSKTAEDLEFKTRALDMFGYTQO	226
	:		:

OY	259	SHMOJFF-KGNRPVRVHDYOEIFLTQPVYDLPMDGKVMIMGMAVHRSNESAKLSR	317
		: : : : : : :	
Dd	227	SDMOIYNQGRKSAPFRNDDYKPEIFLTPQVKADLPFGGRLRMIGAGFVQSNQSGPESR	286
OY	318	SWNRAYLAGMEWKNLTVPMPRIWGHFEKGSSQPDPNDPLDYGYGDVRFPLYLENKS	377
		: : : : : :	
Dd	287	SWNRIFYAAMAGMGKLTGYIPRVYWAFDO-SGDK-NDNPDIADMYMGCVKKLYLRINDRO	344
OY	378	NISSGVRRNPRSGKALODLVYPILGGKGSYGFOIFOGGOSLIQVNHETSRCVGLMUN	437
		: : : : : : : : : :	
Dd	345	NYSVLKRNPKTGYATIEATYFPFKGKLGVYRGFHGESLDINHKONGSIGISLMEN	404
OY	438	DWMGL	442
		:	
Dd	405	DUDGI	409
RESULT	3		
ID	09CL22	PRELIMINARY;	PRT; 306 AA.
AC	09CL22:		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)		
DE	HYPOTHETICAL PROTEIN PM1426.		
GN	PM1426.		
OC	Pasteurella multocida.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Pasteurella.		
OX	NCBI_TaxID=747;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PM70;		
RX	MEDLINE=21145866; PubMed=11248100;		
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;		
RT	"Complete genomic sequence of Pasteurella multocida Pm70."		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3465(2001).		
DR	EMLL; AE006179; AAK03510.1; "		
DR	InterPro: IPR003187; PLA1.		
FW	Pfam: PF02253; PLA1. 1.		
KW	Hypothetical protein; Complete proteome.		
SQ	SEQUENCE 306 AA; 35580 MW; EAF3DBEGCIC22B26E CRC64;		

[illegible]

DE PHOSPHOLIPASE A (EC 3.1.1.32).
GN PLDA OR C11351.
OS Campylobacter jejuni.
OC Bacteriella; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_taxid:197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTCC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jegerl K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail K., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
LR EMBL; AL139078; CAB73778.1; -.
DR Interpro; IPR003187; PLAI.
DR Pfam; PF02253; PLAI; 1.
DR PRINTS; PR01486; PHPHLIPASEA1.
KM Complete proteome.
SQ SEQUENCE 329 AA; 38880 MW; A32731E2B751AC44 CRC64;

Query Match	15.5%	Score 366;	DB 2;	Length 329;
Local Similarity	33.2%	Pred. No. 8.2e-24;		
Matches 105; Conservative	48;	Mismatches 125;	Indels 38;	Gaps 13

```

Oy 130 IKGPQVAVOZETPPIFLMGNEKMLTKRAKOLEYAAKOTPTLSFEEDRRNT--PLW 187
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 43 LKNSVLSIOEONNSSOATOTONSITTKKEKO-----DPSRLALNLYGSESNPLG 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 188 SSRPHNEMVYLPIFMHCKPNKSPPTPSHEARQEPFNEFRAPELKQYSVKKAEDJMG 247
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 97 IS-SYKNMYFLP-FAYSFNSLGVNKKSEA-----KFQLSYVKRLFENLGL 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 248 DSDLMFGYTOOSHWOIFGNKNSRPRVHDYOPELFTLOPVY-SDLPMDGKVRIMGAVH 306
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 DEKYIYAVTQOSMWQIY--EHSSPFRFRTNQPEFIDPLYLKAYEEFFNNLR---VEI 196
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 307 HSGCESAK--LSRNSRWAYLMAGHEMKNLVMPRIWGRIFKESGSGPDDNPDIIDYGY 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 ESNCKSGDENLQSRSMNRITVSTALLYNKFLEVPRLWRI---PENKKODDNPAILHYGN 253
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 365 GDVAFYLQLEKNSISGTVRINPR--SGKAPLDLDIYPL-GKISGTFQIFQSGSLI 422
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 254 FDVNLAY-LGDGYEINILRNRLNFKNHNGKAIQYDGLGYDFNNGIYUWYFNGYGESLI 312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 422 DYNHEATSPFGYGLMN 437
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 DYNKHILQRLSTGLTIS 328
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT	5
Q9SID7	
ID	Q9SID7; PRELIMINARY; PRT; 292 AA.
AC	Q9SID7;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	PHOSPHOLIPASE A.
GN	PLA ₂ .
OS	Yersinia pseudotuberculosis.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC	Yersinia.
NCBI_TaxID=633;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-YPIII PIBI:
RA	Korlyshev A.V., Winzeler E.A., Williams K.J., Oyston P.C.,
RA	Tillett R.W., Wren B.W.;

RT "Biochip-based Signature-Tagged Mutagenesis: Identification and
RT Characterization of *Y. pseudotuberculosis* gene pda essential for
RT virulence in mice."
RT Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ245393; CAB5586.1; --
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
DR PRINTS: PR01486; PROSITE: PS00107.
DR SEQUENCE 292 AA; 33758 MW; 8E712D90AC6BA5 CRC64;

Query Match	15.3%	Score 360.5	DB 2	Length 292
Best Local Similarly	39.2%	Pred. No. 2.1e-23		
Matches 83	Conservative 34	Mismatches 86	Indels 9	Gaps 5

QY 229 ELIROYSVKVKAAADMDINGTSDSLMFGYTOOSHQIFNGKNSRPRVADYDPEIFLGPV- 287
86 EVKFOILSTAFPIWMRIGIGADNSLLASTQKRWMAOANSSESSPRETYNEPOLFELANSTD 145
QY 288 YSDPMDGKVMIGMGAVHNHNGESAKLSRSRNBAYUIMAGMEKNLTVMBRIMGRIPEK 347
Db 146 YELAGW--TFEEVEFGFNGHOSNGKADPTSRMNNAVYTRVAAQKRNLEIDLKPPWRIFESD 203
QY 348 SGSPDPDNDLIDYYGVDVREFYQL-ENNSNIGTVYVNRPSGKALQIDYVYPLKGI 406
Db 204 S---KDNBPDIITKMYGYRLKVGALDSDVSLDG--RYNMNTGYGGAEMGMSYPIIRHV 258
QY 407 SGYFOIQFGYQSLLIDVNHETSRGCVGLMLND 438
Db 259 RPYTQVFSGYESMIDVNFQTRGQVYGLMLND 290

RESULT	6			
Q9L6N9		PRELIMINARY;	PRT;	289 AA.
AC Q9L6N9				
DT 01-OCT-2000	(TEMBLrel. 15, Created)			
DT 01-OCT-2000	(TEMBLrel. 15, last sequence update)			
DT 01-JUN-2001	(TEMBLrel. 17, last annotation update)			
DE PLDA PROTEIN.				
GN PLDA.				
OS Salmonella typhimurium fH2.				
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC Salmonella.				
OX NCBI_TaxID=99287;				
RN [1]				
RP SEQUENCE FROM N.A.				
RC STRAIN-SGSC1412;				
RA washu.				
RT "The Salmonella typhimurium Genome Sequencing Project.";				
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
RN [2]				
RP SEQUENCE FROM N.A.				
RC STRAIN-SGSC1412;				
RA Waterston R.;				
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
DR EMBL: AF233324; AAC33435.1; -				
DR InterPro: IPR003187; PLAL.				
DR Pfam: PF02253; PLAL; 1.				
Q0 SEQUENCE 289 AA; 32967 MW; DA97F5E1651C49C6 CRC64;				

Query Match	15.3%	Score 360;	DB 2;	Length 289;
Best Local Similarly	39.7%	Pred. No. 2.3e-23;		
Matches 85; Conservative	29;	Mismatches 90;	Indels 10;	Gaps 4

[illegible]

```

QY 345 KEGSGSQPDNDPILDYGYGVRLYOLENKSNSIGTVRYNPRSGKALQDLYVPLGK 404
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 200 -----GSTDNDPDTFTKMGYQKLTGYHL-GEAVLSAKQYMMNTGYGAEVGLSYPTK 253
QY 405 GISGYFOIFOGYGOSLIDYNHEATSGVGLMLND 438
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 254 HVRLYQVTSGETESLIDYNFQTRVGVGLMLND 287

RESULT 7
Q924N8 PRELIMINARY: PRT: 292 AA.
ID 0924N8:
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (EC 3.1.1.32).
GN PLDA.
OS Enterobacter agglomerans.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G., Brinkman E., van Boxtel R., Bekkers A.C., Verheij H.M.,
RA Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.";
RT J. Bacteriol. 176:861-870(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120145; PubMed=9921577;
RA Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tommassen J.;
RT "Sequence comparison of outer membrane phospholipases A: implications
RT for structure and for the catalytic mechanism.";
RT Res. Microbiol. 149:703-710(1998).
DR EMBL: AF034414; AAD03498.1; -.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
KW Signal: Hydrolase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 292 OUTER MEMBRANE PHOSPHOLIPASE A.
SQ SEQUENCE 292 AA; 33719 MW; B7516D093B2BEEA CRC64;

Query Match 15.1%; Score 356.5; DB 2: Length 292;
Best Local Similarity 40.1%; Pred. No. 4.7e-23;
Matches 85; Conservative 31; Mismatches 87; Indels 9; Gaps 5;
QY 229 ELKFOVSVKKAEDLMGTDSULMFGYTQOSHMOIFNGKNSRPFVHDYQPEITLQPV- 287
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 86 EVKFOIStGFPIWRIAGTADNSLLGASVYQSRSMWQASNSDESSPFRFTNEPQIFLMAWD 145
QY 288 YSDLPMDGVVRMIGMGAVHNSGESAKLSRSNNRAYLAAGMEKNTLTVPRIRIGTFKGC 347
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 146 YELAGV--TFREVEFGYNHQSNGKADPTSRSDVRYTRLMAQRNLEIDLKFWYRIPESD 203
QY 348 SOSOPDNDPDIIDYGYGVRLYOL-ENKSNSIGTVRYNPRSGKALQDLYVYPLGKGI 406
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 204 S---KDDNDPDKMKYGLKGYALGESVFSDG--RTNMNTGTGAGMGWSYPTTKHV 258
QY 407 SGYFOIFOGYGOSLIDYNHEATSGVGLMLND 438
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 259 RFTYQVFSGYSGSMIDYNFRQTRVGVGLMLND 290

RESULT 8
Q92349 PRELIMINARY: PRT: 297 AA.
ID 092349:
AC 032349:
AC 032349:

```

```

DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Campylobacter coli.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=195;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-UA585;
RX MEDLINE=97230284; PubMed=9119448;
RA Grant K.A., Belandier I., Dekker N., Richardson P.T., Park S.F.;
RT "Molecular characterization of plda, the structural gene for a
RT phospholipase A from Campylobacter coli, and its contribution to cell-
RT associated hemolysis.";
RT Infect. Immun. 65:1172-1180(1997).
DR EMBL: Y11031; CA871915.1; -.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
SQ SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;

Query Match 14.5%; Score 342.5; DB 2: Length 297;
Best Local Similarity 32.4%; Pred. No. 8e-22;
Matches 97; Conservative 37; Mismatches 108; Indels 57; Gaps 11;
QY 178 DLDNRNTPLMSSRPNNPVLPFPMHGKRNRPNT-----P 213
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 16 DLKENNASLSIKKHETQ-----NTQKTPSTKEDFSRIALANYLGESSFNPGLIS 66
QY 214 SHEARQFP-----NEFRAPELKFGVSVKVAEEDLMGTDSULMFGYTQOSHMOIF 264
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 67 SYKMVYFLFPAVSFGSLGGENKRTKMKFQLSIKKRLFEDLLGGEKYYGVYQTSMWQ-- 124
QY 265 NKNNSRPFVHDYQPEITLQPV-YSDLPMDGVVRMIGMGAVHNSGESAK--LSRSNMR 321
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 125 NYKHSSPFRETNYQPEFFDIDLHEDYKFLNNLR--VGLIHESGKGDENLESRSNMR 181
QY 322 AYLMGMEKNTLTVPRIRIGRIFKEGSGOPDNDPDIIDYGYGVRLYOLENKSNSIG 381
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 182 IYASSFYLORLFVPRIMYRI---PENSEDDNPEITHTMGNFIDN-IGSLGNDYFIHL 237
QY 382 TVRYNP--RSGKALQDLYVPL-GKGISGYFOIFOGYGOSLIDYNHEATSGVGLMLN 437
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 238 MLRNNLDFHDKGAVQVDIGYDIFDNGIYWIQYFNGYGSLLIDYNNKRLRSLTAFLIS 296

RESULT 9
Q925241 PRELIMINARY: PRT: 355 AA.
ID 0925241:
AC 025241:
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).
GN HP0499.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalam H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

```


[illegible]

RESULT	12		
Q9F216			
ID	Q9F216	PRELIMINARY;	PRT; 824 AA.
AC	Q9F216;		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	CELLULOSE.		
OS	Bacillus sp.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
OX	NCBI_TaxID=1409;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=KSM-S23/;		
RX	MEDLINE=1036886; PubMed=11193393;		
RA	Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,		
RA	Kobayashi T., Ito S.;		
RT	"Reduced amino acid sequence and possible catalytic residues of a		
RT	thermostable, alkaline cellulase from an alkaliphilic Bacillus		
RT	strain".;		
RL	Biosci. Biotechnol. Biochem. 64:2281-2289,(2000).		
DR	EMBL; AB018420; BAB19360.1; "		
DR	InterPro: IPR001547; Glyco_hydro.F5.		
DR	Pfam: PF00150; cellulase.1.		
DR	PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN.1.		
Q0	SEQUENCE 824 AA; 91564 MW; 655FA940FE1D729B9 CRC64;		

```

Query Match          4.6%: Score 109.5: DB 2: Length 824;
Best Local Similarity 19.7%: Pred. No.0.76; 146; Indels 167; Gaps 27;
Matches 91; Conservative 57; Mismatches 146; Indels 167; Gaps 27;

QY 30 PVAFVDEYRSKND-----LGODNELLIGVQSGASTDTANPL-DEHEPELY-----TT 78
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 PAALAAEGENTRDNPKHLLGNDNVKRPSEAGALQLOEVGOMTLYDQHEKTIQLRMSTH 84
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 ALENTMILNCSAL-----NODIMRLACYPTLVHGE-----PVAVIKTK-----RSI 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 GLOWPEELLINNAVYKALSNMDMSNMIRLAMY-----VGEAGYATNPELLIKQVITDGIELAI 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 RLDEIT---WQGT-----IKGPKVYVQETTD-----IF 146
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 ENDWYIYIDMVIHVAFGDRDPYVAGAKDFRRIALYLRPNPIITIELANEPSSNNNGAG 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 147 LMGNEKGMILTKKDAKQLEAYAAKQTFPLSLSPDLRN-----NPLWSSRP-----H 192
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	201	IPNDEGKAVK-----	EYADPIYEMLRKSGNADNIIIGSPWMSGRPLADNP	LDH	255
Qy	193	NPMTVLPFIEMHGKPNRSPNTPSHBARQFTNEFRAPELK----	FQYSYKVKAAEDL	MGT	247
Db	256	HTMTVTHYEYTSASHAESTESPSE-----	TPNSEGAVNSNTRALENGAVAFATE	-	WGT 308
Qy	248	-----DADLWFGYTOQ-----	SHMOIFENGNS-----	RPRRV-----	274
Db	309	SOASGDGCPFTDEADVMTIEFLNNISGMWMSLTLN-KNEVSGAFTPELCKSNATNLDPG	367		
Qy	275	---HDYQP-ELEFLT-----	-QPV-----YSDLFW-----	DGKVRMIMGAVHHSNGE	311
Db	368	PDHWAPPEELSLTSGEYVRARIKGVNPEIDRTKTKYVLMDPNDGTQ--	GFGVNSDSPNK	425	
Qy	312	SAKLSRSNNRAYLMAGME-----	WKNLVYPMPIRGR	342	
Db	426	ELIADVDENNLTAKVSGLDVSDVSDGFMFNANARISANGWCK	466		

```

RESULT_13
059241
ID 059241 PRELIMINARY; PRT; 821 AA.
AC 059241;
DT 01-NOV-1996 (TREMblrel_01, Created)
DT 01-NOV-1996 (TREMblrel_01, last sequence update)
DT 01-JUN-2001 (TREMblrel_17, last annotation update)
DE ENO-1,4-BETA-GLUCONASE (EC 3.2.1.4) (CELLULOSE) (ENDOGLUCONASE
DE (CARBOXYMETHYL CELLULOSE) .
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_Taxid=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KSM-64;
RA Sumitomo N., Ozaki K., Ito S.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC EMBL: M84963; AAA73189.1; .
DR HSSP; O85465; 1A3H.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase.
DR PROSITE, PS00659; GLYCOSYL_HYDROL_F5; UNKNOWNM.1.
DR HYDROLASE; glycosidase.
SQ SEQUENCE 821 AA; 90910 MW; 73D438FEF0B4DB5C CRC64;

```

Query Match	4.6%;	Score 109;	DB 2;	Length 821;
Best Local Similarity	19.7%;	Pred. No. 0.83;		
Matches	91;	Conservative	51;	Mismatches 150; Indels 170; Gaps 24.
QY	30	PVAFVDEVRSKND-----	LGODNELLGVSQTOSTASTPLNL-DEHEPELY-----	TT 78
DB	24	PTALAAEGNTEDENFKHLIGNDNVKRPEDAGALQLQOEVDGQMTLVDDQGEKTLQIKMSTH		83
QY	79	ALENTMLINCSAL-----	NODIMRLACY-DTLVHGEPFAVIKTKRSIRIDETI----	126
DB	84	GLQWPELLNDNAYKALANDMESNMIRLAMYGEENGASNPPLISRYIKGIDLAIENDM		143
QY	127	-----WQT-----	IKKPOVYVYQETTDP-----	IFLMGN 150
DB	144	YVIVDMHNHAAEDRPDYAGAEDEFRDIAALYPPNNPHILYLAENBPSSNNNGAGAPINN		203
QY	151	EKGMLTKKDAQOLEFAAKQFTPLSLSPDLDRN----	NPLMSSRP-----	HNPMY 196
DB	204	EEGMNAYK-----	ETADPIVELMRDSGNADDIIITVGSPPMSQRPDLAADNPIDIDHTMY	258
QY	197	VLPIF--MHGKPNRS--	PNTPSHEARQFTPNPEFARPELKFQVSUYVYKAAEDMGT-----	247
DB	259	TVHYYTGSMASTESYPPETPNSENGWMSNTRYA-----	LENGAVAFATE--WGTSQANG	312
QY	248	-----DSDLMFGYTQO-----	SHMQLFNGKNS-----	RPRVH----- 275

```

Db 313 DGGRPEADAVNVEFLFNENNISMANSLTN-KNEVSGATFPFELGKSNATSLDGPQVW 371
QY 276 -----DYQPEIFLPQVSDLPW---DGKVRIMGKGAHHNSGES- 312
Db 372 VPEELSLGSEYVRARIKGVNVEP---IDRTKRYTKVLMDFNDGTFKQGGV-----NCDSP 422
QY 313 -----AKLSRNRKAVLIMAGMEKKNLTVMPIRNGR 342
Db 423 VEDVVIENAGALKLGLDASNDVSEGNVWAMARLSADGWCK 464

```

RESULT 14

```

063485 PRELIMINARY; PRT; 602 AA.
ID 063485;
AC 063485;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE C-RAF ACTIVATED ONCOGENE FUSION PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=87172791; PubMed=3550433; Sugimura T.;
RA Ishikawa F., Takaku F., Nagao M.,
RT "Rat c-raf oncogene activation by a rearrangement that produces a
RT fused protein."
RC 1. Similarity: TO THE SER/THR FAMILY OF PROTEIN KINASES.
CL 1. Cell. Biol. 7:1226-1232(1987).
DR EMBL: M15428; AAAA2002.1;
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00221; STRKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Oncogene; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 602 AA; 69209 MW; 88B7BFA90FFB02AC CRC64;

```

Query Match 4.5%; Score 106.5; DB 11; Length 602;
 Best Local Similarity 19.0%; Pred. NO. 0.88;
 Matches 92; Conservative 74; Mismatches 172; Indels 145; Gaps 24;

```

QY 23 QAAQVPRVAFVDEVRSKNDLGODNELLIGVQATQASDTNPRLDENHEPELTALLEN 82
Db 81 ELELNAQVAVLETKKELETAODRL--GIQSQFTRAKEE---LEAEKDLRT---N 131
QY 83 KTMILNCSANMODIMRLACYDVLNGEFTPAVTKRSIRIDETIMQ--TIKSGQVAVQE 140
Db 132 ERLSQEVEYLEDVKRL--NEKLKESNT---TKGELDKLDELQASDYTVVRRKRLQE 186
QY 141 -----TTDPFIKMGNEKG-----MLTKDAKOLEYAAKQFPLS 174
Db 187 KELLNQNSMLNTEKLTDELLALGREGNEILELKCTLEKKEEDAIRSHSESASPSSA 246
QY 175 LSPFLDNNTPLMSRRHNHYVLPYIFMHCGRNRPMTPSHARQFTPNERRAELEKFOY 234
Db 247 LS--SSPNKLSPTGWSQKTP-----VPAOREPAPSGTQEKKINRRGORDSSYYWEI 298
QY 235 SVVKAADLMGTDSDLMFGYTOOSH-----QIFNGKSNRPVHYDOEI----- 281
Db 299 EASEVMSLSTRIGSS---FGTVYKGMKNGVAAVKILKVVDPRTPEQLDAFRREVALKTR 355
QY 282 -----FLTO-----PVYSDL--PMDGKVRMI-----GMGAVNH 307
Db 356 HVNILLFMGYMTKDNLAIVTQWCEGSSLYKHLHVOETKRFQMFOLIDIAQTQAQMDYLHA 415

```

```

QY 308 SNGESAKLSRSMNRKAVLIMAGMEK-----NLTVMPRIKRIKFGSGSO----- 351
Db 416 KNIIHRDMKS--KNIFLHEGLTVKIGDFGLATVSRM-----SGSQVQEPPTGSYLM 466
QY 352 -----PDDNPDLI--DYCYGDVREPLYOLENKSNIQTVRYNPRSGKALQDYYVPL 402
Db 467 APEYIRMQDNNPFSFQSDVYSYGIV--LYEL-----MTGELPYSHINNRDI---IFMV 515
QY 403 GKG 405
Db 516 GRG 518

```

RESULT 15

```

090N32 PRELIMINARY; PRT; 798 AA.
ID 090N32;
AC 090N32;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE PPAR GAMMA COACTIVATOR-1.
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Larroy D., Vidal H., Andreev I. F., Laville M., Langin D.;
RT "Cloning and mRNA tissue distribution of human PPARgamma coactivator-
RT 1."
RC Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF159714; AAD51615.1;
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; Rrm; 1.
DR PROSITE: PS00102; RRM; 1.
DR SMART: SM00360; RRM; 1.
SQ SEQUENCE 798 AA; 91056 MW; F68F9768BD94ELP4 CRC64;

```

Query Match 4.5%; Score 106; DB 4; Length 798;
 Best Local Similarity 24.1%; Pred. NO. 1.5;
 Matches 93; Conservative 36; Mismatches 115; Indels 142; Gaps 27;

```

QY 3 VSLSTLTSLTPCAAILAIQQAQVPRVAFVDEVRSKNDLGODNELLIGVQATQASAST 62
Db 89 LAVLETIDSLPV-----DEGLPSFDALTD-----GDVTIDNE-----ASPSMP 129
QY 63 D-TANPLDENHEPELYTTPAL--ENKTMILN-CSALNODIMRLACYDVLVHGE-----T 110
Db 130 DGTTPRQGAERPSLLKILLAPANTQGLSYNCSGLS-----PQNHANHHRIKTN 179
QY 111 PAVIKTKRSIRIDETIMQ-----IKGPRQ-----VYQETTD-----PIFKMGNE 151
Db 180 PAIVKTTENS-----WSNKAKSICQOQKPORPCSELKLYLTNDNDPRHTKPRENRSS 232
QY 152 KGMILT--KKDAKOLEYAAKQFTPLSPDLDRNNTPLMSSRPHNPW-----YVLP 199
Db 233 RDKCTSKKSHSTQSOQHLQAKPTLSL-----PLTPESPNDKSGSPFNKTIETRLS 285
QY 200 IFMHCGRNRS--PNTPSHARQFTPNERRA--PELKFQV-----SVVKAADLMGT--- 247
Db 286 VELGTGAGLPTTPRPHKANG--DNPRFASPKLSSCTVVPSPSKKPRYSSES--GTQGN 342
QY 248 -----DSDLMFGYTOOSHQIFNG-----KNSRP---FRVNDY-----QPEIFL- 283
Db 343 NSTKKGPQESL---YQQLSKSSVLTGGEHERKTRKPSLRFGDHDYQOSINSKTEILIN 399
QY 284 -----TQPVYSDLPMDKV 297
Db 400 ISQELQDSRQLENKNDVSD--WQGOI 423

```

Fri Nov 30 14:24:21 2001

us-09-787-083-4.rspt

Page 8

Search completed: November 30, 2001, 14:26:27
Job time: 564 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:27:00 ; Search time 28.02 Seconds

(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLPCFAILA.....YNHEATSPGVGLMLNDMNGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	15.8	286	1	PAL_KLEPN
2	358	15.2	289	1	P37446 Klebsiella
3	357	15.1	289	1	P37442 salmonella
4	321.5	13.6	289	1	PAL_ECOLI
5	109	4.6	800	1	PAL_PROVU
6	94.5	4.0	397	1	GUN_BACSL
7	93	4.0	4273	1	GUN_PAEPO
8	92.5	3.9	849	1	PKSM_BACSU
9	92	3.9	324	1	SRKE_BRAOL
10	92	3.9	696	1	YL70_ARCFU
11	92	3.9	34	1	SCK1_SCHPO
12	91.5	3.9	988	1	YJHO_YEAST
13	91	3.9	969	1	SACH_STRSL
14	91	3.9	496	1	CATV_DICDI
15	90.5	3.8	1474	1	AZMG_HUMAN
16	90	3.8	467	1	INVO_MOUSE
17	90	3.8	1158	1	RI14_HUMAN
18	89.5	3.8	1788	1	YP72_CABEL
19	89.5	3.8	486	1	VG14_BPMLS
20	89.5	3.8	669	1	AVI_ALTHA
21	89.5	3.8	1478	1	BCK1_YEAST
22	89	3.8	1840	1	SUIS_RAT
23	89	3.8	1426	1	CURT_MOUSE
24	88.5	3.8	825	1	GUN1_BACSA
25	88.5	3.8	992	1	PHS2_DICDI
26	88	3.7	810	1	ADAM_YEAST
27	88	3.7	1115	1	DP3A_BACSU
28	87.5	3.7	1382	1	MET_RAT
29	87.5	3.7	844	1	PAC_KLUCI
30	87	3.7	1379	1	MET_MOUSE
31	87	3.7	503	1	CD44_RAT
32	87	3.7	666	1	PD14_RAT
33	87	3.7	842	1	LPFC_SALTY
33	87	3.7	882	1	HSSI_RAT

34	86	3.6	882	1	HSSI_HUMAN	P52848 homo sapien
35	86	3.6	1087	1	XYNX_CLOTM	P38535 Clostridium
36	85.5	3.6	353	1	DCUP_BACSU	P32395 bacillus su
37	85.5	3.6	790	1	SEIL_MOUSE	O92496 mus musculu
38	85.5	3.6	794	1	SEIL_HUMAN	O9ub92 homo sapien
39	85	3.6	403	1	P37_MYCHR	P15363 mycoplasma
40	85	3.6	491	1	TY3H_PHASP	P11982 phasianidae
41	85	3.6	560	1	DTXH_CORBE	P00589 corynephae
42	85	3.6	828	1	BGAL_BRAOL	P49676 brassica ol
43	85	3.6	1184	1	ALAC_ARATH	P57792 arabidopsis
44	85	3.6	2231	1	SENI_YEAST	O00416 saccharomyc
45	84.5	3.6	478	1	DHGB_ACTICA	P13650 actinobact

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	286 AA.
PAL_KLEPN				
ID	P37446:			
AC	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PIDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
OX	NCBI_TaxID=573;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94131966; PubMed=8300539;			
RA	Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.;			
RT	"Molecular characterization of enterobacterial pida genes encoding			
RT	outer membrane phospholipase A.";			
RL	J. Bacteriol. 176:861-870(1994).			
CC	- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL: X76901; CAA54223.1; -			
DR	PIR: B36971; B36971.			
DR	PIR: S40129; S40129.			
DR	InterPro: IPR003187; PLA1.			
DR	Pfam: PF02253; PLA1; 1.			
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.			
FT	CHAIN	1	20	BY SIMILARITY.
FT	ACT SITE	21	286	PHOSPHOLIPASE A1.
FT	ACT SITE	161	161	BY SIMILARITY.
SO	SEQUENCE	286 AA;	3639P863085108A3	CRC64;

Query Match 15.88; Score 373; DB 1; Length 286;
Best local Similarity 42.18; Pred. No. 3.3e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

	Matches	84: Conservative	30: Mismatches	90: Indels	10: Gaps
QY	226	RAPELKEFVSVKAAEDLWGTSDLMFGYTOQSHWQIFGNKSNRPVRHYDQPEILFQ	285		
		: : : : : : : : : : : : : : : : : : : : : : : : :			
Db	83	RKDEKFPQLSLAFPLMRGILGPNNSVLGASYQKQSMWQLSNKESSPPRETTYEPQLFGF	142		
QY	286	PV-VSDLPMDCKVMNIGMGAHVHNSGSAKLSRSMNRNRYILMAGEMKNTLTMPRIWGIIF	344		
		: : : : : : : : : : : : : : : : : : : : : : : : :			
Db	143	ATDRFPAGW-TLIRVEVGGYVHNSGRSDPTSRWNRLTYRLMAENGNWLVKRPVYI-	199		
QY	345	KEGSGOSPPDDPDILDDYGYGDVRFYQLEKNSIGSTVRNPPSPGALOLDVVPYLGK	404		
		: : : : : : : : : : : : : : : : : : : : : : : : :			
Db	240	-----GSTDNDMPDITKVMGYQLKIGVHL-GEAVLSAKGQVNMWTVGGGAIVGLSTPYTK	253		
QY	405	GISGTFQIFGQYGGOSLDIDYVHEATSPFGVGLMD	438		
		: : : : : : : : : : : : : : : : : : : : : : : : :			
Db	254	HVRITYGYSGESGLIDYNFNQTRVGVGLMD	287		
RESULT	3				
PAL-ECOLI	STANDARD:	PRT:	289	AA.	
ID	PAL-ECOLI				
AC	P00631;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DE	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.33) (DETERGENT-RESISTANT				
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-				
DE	ACTLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE) (OM PLA).				
GN	PLA OR B3821 OR Z5342 OR ECS4751.				
OS	Escherichia coli, and				
OS	Escherichia coli 0157:H7.				
OC	Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Escherichia				
OX	NCBI_Taxid=362, 83334;				
UN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=81517492; PubMed=6397464;				
RA	Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,				
RA	Iroue K., Ikeda H., Setiguchi M., Nojima S.;				
RT	"The DNA sequence encoding pldA gene, the structural gene for				
RT	de detergent-resistant phospholipase A of E. coli.";				
RL	J. Biochem. 96:1655-1664(1984).				
UN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-K12 / MG1655;				
RC	MEDLINE=92358234; PubMed=1379743;				
RA	Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;				
RT	"Analysis of the Escherichia coli genome: DNA sequence of the region				
RT	from 84.5 to 86.5 minutes.";				
RL	Science 257:771-778(1992).				
UN	[3]				
RP	REVISION TO 14-15.				
RC	STRAIN-K12 / MG1655;				
RC	MEDLINE=97426617; PubMed=9278503;				
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,				
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,				
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,				
RA	Ma B., Shao Y.;				
RT	"The complete genome sequence of Escherichia coli K-12.";				
RL	Science 277:1453-1474(1997).				
UN	[4]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-0157:H7 / EDL933 / ATCC 700927;				
RC	MEDLINE=21074935; PubMed=11206551;				
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,				
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,				
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller E.,				
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,				
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,				
RA	Weich R.A., Blattner F.R.;				
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";				

RL Nature 409:529-533(2001).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RX MEDLINE-21156231; PubMed-1258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RT Complete genome sequence of enterohemorrhagic *Escherichia coli*
 O157:H7 and genomic comparison with a laboratory strain K-12.*;
 RL DNA Res. 8:11-22(2001).
 RN [6]
 RP SEQUENCE OF 30-289 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-83003590; PubMed-6383820;
 RA de Geus P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
 RT "The pro- and mature forms of the E. coli K-12 outer membrane
 phospholipase A are identical.";
 RL EMBO J. 3:1799-1802(1984).
 RN [7]
 RP SEQUENCE OF 174-289 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-87115164; PubMed-3027506;
 RA Irino N., Nakayama K., Nakayama H.;
 RT "The recQ gene of *Escherichia coli* K12: primary structure and
 evidence for SOS regulation.";
 RL Mol. Gen. Genet. 205:298-304(1986).
 RN [8]
 RP MUTAGENESIS OF SER-172.
 RX MEDLINE-94131966; PubMed-8300539;
 RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
 Verheij H.M., Tomassen J.;
 RT "Molecular characterization of enterobacterial pldA genes encoding
 outer membrane phospholipase A.";
 RL J. Bacteriol. 176:861-870(1994).
 RN [9]
 RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
 RX MEDLINE-91249806; PubMed-2040286;
 RA Horrevorts A.J.G., Verheij H.M., de Haas G.H.;
 RT "Inactivation of *Escherichia coli* outer-membrane phospholipase A by
 the affinity label hexadecanesulfonyl fluoride. Evidence for an
 active-site serine.";
 RL Eur. J. Biochem. 198:247-253(1991).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
 CC DORMANT IN NORMAL GROWING CELLS.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
 CC LYSIS OR TEMPERATURE SHOCK.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X02143; AAA26081.1; -
 CC EMBL: M87049; AAA67617.1; -
 CC EMBL: AE000458; AAC76824.1; -
 CC EMBL: AE005613; AAG59017.1; -
 CC EMBL: AP002567; BAB38174.1; -
 CC EMBL: M30198; AAA24516.1; -
 CC PIR: A00771; PSECA.

DR PIR: A22133; PSECA.
 DR PIR: S30711; S30711.
 DR Ecogene: EG10738; PldA.
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
 KM Complete proteome.
 FT SIGNAL 1 20
 FT CHAIN 21 289 PHOSPHOLIPASE A1.
 FT ACT_SITE 164 164
 FT MUTAGEN 172 172 S->F: INACTIVE PROTEIN.
 FT CONFLICT 14 15 LP -> FA (IN REF. 2).
 FT CONFLICT 30 33 DAPA -> WTRQ (IN REF. 6).
 SQ SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CRC64;
 Query Match 15.1%; Score 357; DB 1; Length 289;
 Best Local Similarity 39.3%; Pred. No. 6,7e-22;
 Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;
 QY 226 RAPELRFQVSYKKAADLMGTSDFLNFYGTQSHQIFNGKNSRRPRVHDYDFELFQ 285
 DB 83 RKDEVKFQLSLAPLWGLGIPNSVLGASYQKSWQLSNSSESPRETYEPQLGLGF 142
 QY 286 PV-YSDLPMGKVMIGMAVHHSNGESAKLSRSNMNPAVYLMAGEMKRLTYMPRIWGRIF 344
 DB 143 ATDYRFAGW--TLRDVEKGNHDSNGSDPTSRSMNRLYRIAMENGNLVEYKPVIV- 199
 QY 345 KEGSGSQDDNPDLIDYGGVDFLYOLENKSNIQGTVRNPSRGALQLDYVPLGK 404
 DB 200 -----GNTDDNPDTIKYGYQLKIGYHL-GDAVLSAKGQNMWMTGAGELGLSTPYTK 253
 QY 405 GISCYFOIFGCGSLIDYHNEATSRFGVGLND 438
 DB 254 HVRLYTQVYSGYGLIDYFNQTRVGWGLND 287
 RESULT 4
 ID PAL_PROVU STANDARD; PRT; 289 AA.
 AC P37447;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
 DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
 DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
 GN PldA.
 OS Proteus vulgaris.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Proteus.
 OX NCBI_TaxID=585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94131966; PubMed-8300539;
 RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
 Verheij H.M., Tomassen J.;
 RT "Molecular characterization of enterobacterial pldA genes encoding
 RT outer membrane phospholipase A.";
 RL J. Bacteriol. 176:861-870(1994).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X76902; CAA54224.1; -.
DR PIR: C36971; C36971.
DR PIR: S40130; S40130.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KW Hydrolase; lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 289 PHOSPHOLIPASE A1.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 289 AA; 32944 MW; D7516CFEB406997 CRC64;

Query Match 13.6%; Score 321.5; DB 1; Length 289;
Best Local Similarity 33.3%; Pred. No. 5.3e-19;
Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

QY 179 LDRNPTLWSSRPHNPVYLPFMHGKPNRSPNTPSHEARQFTPEFRAPELKFGYSKV 238
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 43 LQEHNP-FTLYPESNLYLTYY-----TSDLNKKALIESYNNSDNA-NKDEYKFOLSLAF 95

QY 239 KAEDLMGDSOLMGFTYQSHMOIFNGKNSRPFVHYDQPELFTQPY-YSDLPMDGV 297
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 96 PLMRGILGNSLSLGASYTORSMWOLSNTEGSAFPRETNPEQLFGFATYDYSVGM--TL 153

QY 298 RMIGCAVHNSGESAKLSRWNRAVYLMAGMEKMLTVPRIMGRIFKSGSSQPDNDP 357
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 154 RAEFGYVHNSGRSDPPTSRSMNRKLSRLMAONGMLVEYKFWYI-----GDTSDNKN 207

QY 358 ILDYGYGDFRFLYOLENKSNIQVRYNPRSGKALQIDVYYPGLKGISGYFOIGYG 417
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 208 IYKMGYQYGLKIGYGL-GEAVLSAKGYWMNMGYGAELGVSYPITKHRYFTQYSGYG 266

QY 418 QSLIDYNEATSFYGLMLND 438
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 ESLIDYDFNQTRVGMVGLND 287

RESULT 5
GUN_BACSI STANDARD: PRT; 800 AA.
AC P06564;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE (ALKALINE CELLULOSE).
OS Bacillus sp. (strain 1139).
OC Bacteria; Firmicutes; Bacillales; Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1411;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87085443; PubMed=3098909;
RA Fukumori F., Kudo T., Narahashi Y., Horikoshi K.;
RT "Molecular cloning and nucleotide sequence of the alkaline cellulase
RT gene from the alkalophilic Bacillus sp. strain 1139."
RL J. Gen. Microbiol. 132:2329-2335(1986).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1- MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES
CC ONE ALKALINE CELLULOSE (PH OPTIMUM 9.0). THIS STRAIN IS NOT A TRUE
CC CELLULOTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO
CC HYDROLYSE NATIVE CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15743; AAA22305.1; -.
DR EMBL: D00066; BAA00045.1; -.
DR PIR: A29003; A29003.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 800 ENDOGLUCANASE.
FT ACT_SITE 190 190 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 305 305 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 800 AA; 88602 MW; 7CCA4D7BBDAD55CF CRC64;

Query Match 4.6%; Score 109; DB 1; Length 800;
Best Local Similarity 19.7%; Pred. No. 0.47;
Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;

QY 30 PVAFVDEVRKND-----LGQDNELLIGVQASATQASDTNAPL-DEHEPELY-----TT 78
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 25 PTALAEAGNTRFEDNFKHLGNDNVRKRPSEAGALQLEVDGQMTLDVQHEKIQLRGMSTH 84

QY 79 ALENKTMILNCSAL-----MODIMRLACY-DTLVHGFTPAVITKRSIRIDETI---- 126
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 GLQWPEELINDAKYALANDMDSNMIRLAMYGENGYASNPLISRYIKGIDLAIENDM 144

QY 127 ----WQT-----IKKRPVYQETTD-----IFLMGN 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 145 YVIVDMVHVAPEDPDPVYAGAEDEFRIALYLPNNPHIYELANBSSNNNGAGIPNN 204

QY 151 EKGMLTKDKAKOLEYAAKQFTPLSLFDLRN-----NPLMSSRP-----HNPMY 196
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 EEGMNAVK-----EYADPIVEMLRDSGNADNIIIVGSSQRPDLADNPIDIDHTTY 259

QY 197 VLPPIF--MHGKPNRS--PNTPSHEARQFTPEFRAPELKFGYSVYKAAEDLMGT-- 247
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 260 TVHFTYTGSHAASTESYPPETPNSERGNVMSNTRYA-----LNGVAVFATE--WGISQANG 313

QY 248 -----DSDLWFGYQQ-----SHMOIFNGKNS-----RPRVH----- 275
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 314 DGGPYFDEADVWIEFLNENNISMANMSLTN-KNEVSGAFTPELKSNAATSLDPGPDQY 372

QY 276 -----DYOPELFTQPYVSDLPW---DQKVMKMGAVHNSGES- 312
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 373 VPEELSLGGEYRARIKGVNERP---IDRTKYTKYLVMPDFNDTKGFGV-----NGDSP 423

QY 313 -----AKLSRSWNRAVYLMAGMEKMLTVMPRIWGR 342
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 424 VEDVYIENAGALKLSGLDASNDVSEGNVMANARLSAGCK 465

RESULT 6
GUN_PAEPO STANDARD: PRT; 397 AA.
AC P23548;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ENDOGLUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULOSE).
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Paenibacillus.
OX NCBI_TaxID=1406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90170877; PubMed=2307659;
RA Baird S.D., Johnson D.A., Seligy V.L.;

```

```

RT      "Molecular cloning, expression, and characterization of
RT      endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
RT      circulans".
RL      J. Bacteriol. 172:1576-1586(1990).
CC      -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC      LINKAGES IN CELLULOSE.
CC      -1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC      HYDROLASES).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: M33791; AAA22631.1; -.
CC      DR      PIR: A35136; A35136.
CC      DR      HSSP: P54583; 1ECE.
CC      DR      InterPro: IPR001547; Glyco_hydro_F5.
CC      DR      Pfam: PF00150; cellulase_1.
CC      DR      PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
CC      KW      Cellulose degradation; Hydrolase; Glycosidase.
CC      FT      ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
CC      FT      ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
CC      SO      SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

Query Match          4.0%; Score 94.5; DB 1; Length 397;
Best Local Similarity 25.7%; Pred. No. 2.7;
Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY      301 GMAVHNSNESAKLSNMRAVILMAGMEKNLTVMRIGR-----IFKSG----- 347
DB      43 GNRIVBSGKEAFNGLNW-----FLETPNNT-LHGLMSRSDMLDQYKEGYMLIR 95
QY      348 -----SGSOPDD-----NPDLIDYGVGVFLYQLEKNSINISGV---RYNPRS 389
DB      96 LPSYNSOLFSSSRDSDIDYHKNPDV---GLNPQIMDKLEKAGQGIQIILDRHRPGS 152
QY      390 GKALQLDIV--YPLGKISGYFQIFQGY-----GOSLIDVNEATSEFGVGLMLNDW 439
DB      153 G-GOSELMYTSQYPERSRMISDMKMLADRYKNPNPTVIGADLHNEBHGQASMGTSMDW 210

RESULT 7
PKSM_BACSU          STANDARD;          PRT; 4273 AA.
AC      PA0872; O31781;
DT      01-FEB-1995 (Rel. 31, Created)
DT      20-AUG-2001 (Rel. 40, Last sequence update)
DE      20-AUG-2001 (Rel. 40, Last annotation update)
DE      PUTATIVE POLYKETIDE SYNTHASE PKSM.
GN      PKSM OR PKSY.
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group.
OC      Bacillus/Staphylococcus group; Bacillus.
OX      NCBL_taxid=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Kunit F., Ogasawara N., Yoshikawa H., Danchin A.;
RL      Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE OF 1-1763 FROM N.A.
RA      STRAIN-168 / PB1424;
RA      Tognoni A., Grandi G.;
RL      Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
CC      THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
CC      SECONDARY METABOLISM.
CC      -1- COFACTOR: CONTAINS 4 COVALENTLY BOUND PHOSPHOPANTETHEINES

```

```

CC      (POTENTIAL).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: Z99113; CAB1603.1; -.
CC      DR      EMBL: Z35133; CAA84505.1; -.
CC      DR      Subtilist; BG10931; PKSM.
CC      DR      InterPro: IPR000794; ketoacyl-synt.
CC      DR      InterPro: IPR001601; Meth-transf.
CC      DR      InterPro: IPR003880; Phosphopant_attach.
CC      DR      InterPro: IPR000051; SAM_bind.
CC      DR      Pfam: PF00550; pp-binding; 1.
CC      DR      PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
CC      DR      PROSITE: PS50075; ACP_DOMAIN; 4.
CC      DR      TRANSFERASE; Acyltransferase; Multifunctional enzyme; Repeat; Complete proteome.
CC      KW      Phosphopantetheine;
CC      FT      DOMAIN 295 364 ACP_CARRIER (ACP) 1.
CC      FT      DOMAIN 396 834 ACP_CARRIER (ACP) 1.
CC      FT      DOMAIN 2190 2258 ACP_CARRIER (ACP) 2.
CC      FT      DOMAIN 3222 2737 BETA-KETOACYL SYNTHASE 2.
CC      FT      DOMAIN 3532 3947 BETA-KETOACYL SYNTHASE 3.
CC      FT      DOMAIN 3410 3483 ACP_CARRIER (ACP) 3.
CC      FT      DOMAIN 4140 4209 ACP_CARRIER (ACP) 4.
CC      FT      BINDING 327 327 PHOSPHOPANTETHEINE (POTENTIAL).
CC      FT      BINDING 2222 2222 PHOSPHOPANTETHEINE (POTENTIAL).
CC      FT      ACT_SITE 2476 2476 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
CC      FT      BINDING 3446 3446 PHOSPHOPANTETHEINE (POTENTIAL).
CC      FT      ACT_SITE 3690 3690 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
CC      FT      BINDING 4172 4172 E -> V (IN REF. 2).
CC      FT      CONFLICT 103 103 PHOSPHOPANTETHEINE (POTENTIAL).
CC      FT      CONFLICT 276 276 O -> E (IN REF. 2).
CC      FT      CONFLICT 289 289 T -> S (IN REF. 2).
CC      SO      SEQUENCE 4273 AA; 477459 MW; 3BBFC1A250AB5A CRC64;

Query Match          3.9%; Score 93; DB 1; Length 4273;
Best Local Similarity 19.2%; Pred. No. 97;
Matches 102; Conservative 79; Mismatches 163; Indels 188; Gaps 29;

QY      14 PCFAIILAQQAQAV-PNPVAFVDE-----VRSKN-----DLGQ 45
DB      3597 PLFQIIPKREASMDPPQPIFLEAMHTFEDAGYMGDRIRKSGGVYVGESEGYAHLTG 3656
QY      46 DNELLIVQASAKTASDTDPANPLDEHEPELYTTALEKTKMLI---KCSALNODIMRLACY 102
DB      3657 DTDYINGTQNTATLSAR--IYVALDLKGPVNAITLTAACSSGVAIHQAQSAALRQG---DCE 3710
QY      103 DTLVHGEPNAVYKTKRSIRLDETITQTIKGPVYQVQETDPPIFMNGEKGMLTKKDAQK 162
DB      3711 MALAAGYTLNI-----SHMSFELTRAEMLSPNGCKVVDQDANG 3750
QY      163 L-----EYAAKQFTPLSLSPDLRNNTPLWSSRPNPVYLPFEMHGPN--RSPNPSHE 216
DB      3751 LVPGEAVALVILKPLSKAIE-DKDH-----YGCIKASGVNVDGKTNGITARN-PFSQ 3801
QY      217 ARQ-----FTREFFAPAELEKQVY-----VKVVAADLMDGTSDLMFGYIQGSHWQ 262
DB      3802 AELLINIEYKKEINPLDIIQYMAHSTGSNLGDPLEVALTSVFSK-----YTKQKQFC 3854
QY      263 IFNGKNSRPFVRHVDYQEIFLTQPVYSDDLPMQGVKRMIGM-----GAVHHSNGES- 312
DB      3855 MIS--SIKPLIGHFFAA-----SGVVALISHLMAKNKIITATHHCSENP 3898
QY      313 -----AKLSRSMRAYLMAGMEKNLTVMRIGRIFKSG-----SGSO----- 351

```

```

Db 3899 YIPKESPVLCENKNSWK-----KN--QKPRM-GTISTTIGSGTNAHVEEYI 3946
QY 352 PDDNPDLITYGYGDV-----RELYOLENKSNIS-GTVRYNPRSGKAL 394
Db 3947 PDDQSTQHQHSGPOLFIYSAONDDRLQDACRMIAVLBQNHLSPDVAAYTLQVGRKAM 4006
QY 395 QL-----DYVYPL-GRGISGYFO-IFGYSQSLDYHNEA 427
Db 4007 EARLAIVANNQOLYAKLEYEAMKNGVSGQQRSLYGTGILEODEA 4058

RESULT 8
SRK6_BRAOL STANDARD; PRT: 849 AA.
ID SRK6_BRAOL STANDARD; PRT: 849 AA.
AC 009092;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
DE (S-RECEPTOR KINASE) (SRK).
CN SRK6.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, S656; TISSUE=Stigma;
RX MEDLINE=92020942; PubMed=1681543;
RA Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
RT "Molecular cloning of a putative receptor protein kinase gene encoded
at the self-incompatibility locus of Brassica oleracea.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
CC -1- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
(FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
CYTOPLASMIC DOMAIN.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN - ADP + A PHOSPHOPROTEIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PETIOLE AND ANOTHER.
CC -1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
B. OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
CC -1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M76647; AAA33000.1; ALT_TERM.
DR HSBP; P11362; IREG.
DR InterPro: IPR001480; B_lectin.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003609; Pan_kinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR InterPro: IPR000858; Slocus_glycop.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00954; S_locus_glycop; 1.
DR SMART; SM00108; B_lectin; 1.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW transferase; Serine/threonine-protein kinase; Signal; ATP-binding;

```

```

KM Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
FT SIGNAL 1 32
FT CHAIN 33 849
FT DOMAIN 33 446
FT TRAMEM 447 466
FT DOMAIN 467 489
FT DOMAIN 528 779
FT NP_BIND 534 542
FT BINDING 556 556
FT ACT_SITE 653 653
FT CARBOHYD 47 47
FT CARBOHYD 120 120
FT CARBOHYD 196 196
FT CARBOHYD 260 260
FT CARBOHYD 314 314
FT CARBOHYD 389 389
FT CARBOHYD 442 442
SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 849;
Best Local Similarity 21.9%; Pred. No. 11;
Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;

QY 105 LVHGETPAVYIKTKRSIRLDETWTQIKKPVQVYQETDPIFMG---NEK---GMLTK 157
Db 24 LTHPALSTYINT-----LSTESTLITSNKTLV---SPGSIFEYGFRRNSHWYGLMWYK 75
QY 158 K-DARQLEVAKQFTPLISFP-----LDRNNTPLSSRPHNMVYLPFIMHCK 205
Db 76 KVSDDRYVAVANRNDPLSNAIGTLKISGNNLVLLDHSNKPVMTN-----LTRGN 125
QY 206 PNRSP-----NTPSHARQPT-----PNEFRAPELKFQVSVYKAAEDLM 245
Db 126 -ERSPVVAVELLANGNVNMDSSNMDASEYLMQSPYPTDTLLPEKKGIVNLT----- 177
QY 246 GTDSDLMEGYTQSHWQIFNGKNRSRFRVHYQ-----PEFLTQ---PVYSDLPMGCK 296
Db 178 GLNRFV-----TSWRSSDDPSGDNF---SYKLTQSLPERYLSRENPMHRSQPMWG- 226
QY 297 VMNMGAVH-----HSGESAKLSRSMNRAY-----LMAGMEKNKLVMP--RI 339
Db 227 IFFSGIPEQKLSYVAVNFENNEVATTFKMTNNSFSRLTIEGFGRLTWTWPSIRI 286
QY 340 WGRIFKSGSGQDDNPDLIDY---YGYGVRELYOLENKSNISGTVR-YNPRS 389
Db 287 WNRFW-----SSPVD-PQDDTYIMCGPYAYCDV-----NTSPVCNCIOGFNPRN 329

RESULT 9
ID YL70_ARCFU STANDARD; PRT: 324 AA.
YL70_ARCFU STANDARD; PRT: 324 AA.
AC 028112;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHEITICAL PROTEIN AF2170.
GN AF2170.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervlavage A.R., Graham D.E., Kyprides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodok A., Zhou L.,

```


QY 313 AKLSFNMNRAVIMAG---MEMKNLTVMPRING---RIKESGSGSPDDNDPILDIYY---G 363
 Db 458 ATTNFCGCTTELEAEVLEEDKGYTKQVDFMSLGVLFEMCCGMSBPYAPDVQMYRNIA 517
 QY 364 YGDVRF---LYOLEKNSINSGTVRYNPRSGKAL 394
 Db 518 FGKVRFPKGVLSSEGRSFVRGLNRPNHRLGAV 551

RESULT 11

YJH0_YEAST STANDARD: PRT: 888 AA.
 ID YJH0_YEAST
 AC P40361.1
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHEETICAL 104.3 KDA PROTEIN IN SMC3-MRPL8 INTERGENIC REGION.
 GN YJ070C OR J1095 OR HRD888.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=95282514; PubMed=7762302;
 RA Vandenbol M., Durand P., Dion C., Portetelle D., Hilger F.;
 RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
 RT Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
 RT L8.";
 RL Yeast 11:57-60(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Sor F.J.;
 RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; 234288; CAAB4052.1; -
 DR EMBL; 249345; CAAB9362.1; -
 DR EMBL; 249344; CAAB9361.1; -
 DR EMBL; X88851; CAAG1309.1; -
 DR PIR; S47120; S47120.
 DR SGD; S0003606; YJL070C.
 DR InterPro: IPR001365; A_deaminase.
 DR Pfam: PF00962; A_deaminase; 1.
 KW Hypothetical protein; Hydrolase.
 SQ SEQUENCE 888 AA; 104263 MW; DB31A8086224114D CRC64;

Query Match 3.9%; Score 92; DB 1; Length 888;
 Best Local Similarity 23.1%; Pred. No. 13;
 Matches 66; Conservative 36; Mismatches 126; Indels 70; Gaps 15;

QY 128 QTIKGPQVVOE---TTDPIFLMGNEKMLTKKDA-----KOLEYAAKQFTPLSL 176
 Db 2 QAVERRPILDEYQNSYTKPNFKKEARVLSNDGVSPLKOKKISVDDMDMISLP 61
 QY 177 FDDLRN---NRPPLW---SSRPINPMVYLPIFMHGKPNR-----SNTTSH-----EAKQF 220
 Db 62 TEPFRQWLVSGPMFEDLEDEENKIDPLPSVSHYNGSDFSVSYTSPNLTGTGETKDL 121
 QY 221 TPNEFRAPLKFQVSVKVKAAEDLMG-----TSDLMFGYTNQ---QSHQIENGKNSRPF 273
 Db 122 FINPF---ELVSQMKRKRIIAASKODGISININDTEKMWLFKPLPKTFKREDDK-----R 173

QY 274 VHDYQPEIFLQPVYSDLPMDGKYRMIGCAVHNSNGESAKLSRSMNRAVIMAGEMKNL 333
 Db 174 FQD-----PDSDLNDOSDSTGCGAATPHRRGYYPPSYFTHHYYTTSGLKAGKN 224
 QY 334 TVMPRIKRIKESGSGSPDDNDPILDIYYGVDVRFYLOLEKNSINSGTVRYNPRSGK 391
 Db 225 IKVP-YTGEYF-----DLEDY-----KKQYIYHLSNQENTONPL--SPYSSK 263

RESULT 12

SACB_STRSL STANDARD: PRT: 969 AA.
 ID SACB_STRSL
 AC Q05242.1
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE LEVANSUCRASE PRECURSOR (EC 2.4.1.10) (BETA-D-FRUCTOFURANOSYL
 DE TRANSFERASE) (SUCROSE 6-FRUCTOSYL TRANSFERASE).
 GN FTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=93322332; PubMed=8331080;
 RA Ratham C., Giffard P.M., Jacques N.A.;
 RT "The cell-bound fructosyltransferase of Streptococcus salivarius: the
 RT cardoxyl terminus specifies attachment in a Streptococcus gordonii
 RT model system."
 RL J. Bacteriol. 175:4520-4527(1993).
 CC -1- CATALYTIC ACTIVITY: SUCROSE + (2,6-BETA-D-FRUCTOSYL)(N) =
 CC GLUCOSE + (2,6-BETA-D-FRUCTOSYL)(N+1) (OTHER SUGARS CAN
 CC ALSO ACT AS FRUCTOSYL ACCEPTORS).
 CC
 CC -1- SUBCELLULAR LOCATION: SECRETED. CELL-WALL BOUND.
 CC
 CC -1- SIMILARITY: BELONGS TO FAMILY 68 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; 108445; AAA71925.1; -
 DR InterPro: IPR003469; Glyco_hydro_68.
 DR Pfam; PF02435; Glyco_hydro_68; 1.
 KW Transferase; Glycosyltransferase; Signal; Cell wall.
 FT SIGNAL 1
 FT CHAIN ? 969 LEVANSUCRASE.
 SQ SEQUENCE 969 AA; 103983 MW; D389B5832ACF735A CRC64;

Query Match 3.9%; Score 91.5; DB 1; Length 969;
 Best Local Similarity 19.9%; Pred. No. 16;
 Matches 85; Conservative 54; Mismatches 161; Indels 127; Gaps 20;

QY 53 VQATQASASTQANPLDHEPELYTTALEKTMILNCALMODIMRLA---CYTTLVHG 108
 Db 155 VEAPTSTASSBA---DTH-TEVDLKVSENSAANANSLKGRKISYEENMTSQIV-A 209
 QY 109 ETPAVIKTKRSRLDETITWQITKG-PQVVOETTPDPIFLMGNEKMLTKKDAKOLEVA 167
 Db 210 LTEETIKALKNVDFSD---DAIKGTISLTYRNLDIV-----ASFLKQDSKLAVPYF 259
 QY 168 KQFTPLSL-SFD-----LDRNNTPLWSSRPHPM-----YVLPFMHGKPNRSP 210
 Db 260 KADTLINNPANNTYDAQYMKKEIDVDMSWVQDAKSGVSNMNGYOLVIMAGAPNNNS 319
 QY 211 NTPSHEARQFTPNNEFRAPLKFQVSVKVKAAEDLMGTDSDLMFGYTOOSHWOIFNGKNSR 270


```

Db 320 NHLYLRKYGDNDF-----THW-----KNAG 341
QY 271 PERVHVOPIEFILTOPYSLPMDGKVRMGKAV--HHSNGESAKISRSQNR---AYLM 325
Db 342 P-----IFGYMLEDQOQMSGATVNSDSIOLYTKNDPSGGKLMWOLASATLN 392
QY 326 AGMEMKLTVMPIRMGRIFKEGSG---SOP-----DDNPDILDYGYGD--- 366
Db 393 LAVENDEVYKSVENDHILFEGGNYHQSTFKEMSTFDDDNHNDGNDPDRDNTCLRPHI 452
QY 367 -----VRFI-----YOLENK-----SNISQTVRNPBSGKALQLDYVYPLGKIS 407
Db 453 IEDNGSRHYLFESNTGDENYQGEKQIYKMSNYGGDAFNLSPLINYNNHLYNLASMAN 512
QY 408 GYFOIFD 414
Db 513 GSIGILK 519

```

RESULT 13

```

CATAL_DICDI STANDARD: PRT; 496 AA.
ID CATA_DICDI
AC 077229;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA OR CAT.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Metazoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Foote C., Alexander H., Alexander S.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) - O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF090443; AAC36743.1; -.
DR HSSP: P00432; 7CAT.
DR DICTYDB: DD02727; CATA.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase; 1.
DR PRINTS: PR00067; CATALASE.
DR PRODOM: PD000510; Catalase; 1.
DR PROSITE: PS00437; CATALASE; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
KM Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 54
FT ACT_SITE 128 BY SIMILARITY.
FT BINDING 338 BY SIMILARITY.
FT SITE 494 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SITE 496 MICROBODY TARGETING SIGNAL (POTENTIAL).
SO SEQUENCE 496 AA; 55683 MW; 683318B8FAFDE6 CRC64;

```

Query Match 3.9%; Score 91; DB 1; Length 496;
 Best local Similarity 23.5%; Pred. No. 7;

```

Matches 56; Conservative 32; Mismatches 94; Indels 56; Gaps 12;
QY 61 STQANPLDEHEPELYTALFNKTNLNGSLNODINRIACDYDTLVGEPAYIKTRSI 120
Db 7 TTSSGSPIDNN-----LNSMTAGVNGPILIDFTLI--DLAHDRE----- 46
QY 121 RLDEETIMQITIKKPOVYVOETTD-----PIF-----LMGNEKML-T 156
Db 47 RIPEPVYHA-KGAGAHGYFEVTSDDYKCKAKFLNKVGRKTFLETFEFSYSGKSSDS 105
QY 157 KDAKOLEYAAKOPTPLSLFDLRNNTPLMSSRPHPMYYLPIFMHGKPNRSPPTSHE 216
Db 106 ERDPR--GFAVKFYTEEG-NEDWVGNTMPVFIFIDPSK---FPDFHTQ-KRNPQTNCKD 158
QY 217 AROFTPNFAPRLKFOYSYKVAEDLMGDSLMGCTQOSHWOLFNGNSPPFY 274
Db 159 PNMFDFLGOTPESTHOVSILFSDR---GTPKSYRRMHGFSHTLKFVNAQGRPYW 212

```

RESULT 14

```

A2MG_HUMAN STANDARD: PRT; 1474 AA.
ID A2MG_HUMAN
AC P01023;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-2-MACROGLOBULIN PRECURSOR (ALPHA-2-M).
GN A2M.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190481; PubMed=2581245;
RA Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hiorns L.R., Fey G.H.;
RA "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and
RA assignment of the chromosomal locus".
RA Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT HIS-704.
RX MEDLINE=92246939; PubMed=1374237;
RA Mathijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,
RA Marynen P.;
RA "Structure of the human alpha-2-macroglobulin gene and its promotor.".
RA Biochem. Biophys. Res. Commun. 184:596-603(1992).
RN [3]
RP SEQUENCE OF 24-1474.
RX MEDLINE=84239807; PubMed=6203908;
RA Sottrop-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RA "Primary structure of human alpha 2-macroglobulin. V. The complete
RA structure."
RA J. Biol. Chem. 259:8318-8327(1984).
RN [4]
RP ERRATUM.
RA Sottrop-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RA J. Biol. Chem. 260:6500-6500(1985).
RN [5]
RP INHIBITORY SITE.
RX MEDLINE=84030513; PubMed=6195065;
RA Virca G.D., Salvesen G.S., Travis J.;
RA "Human neutrophil elastase and cathepsin G cleavage sites in the bait
RA region of alpha 2-macroglobulin. Proposed structural limits of the
RA Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).
RN [6]
RP INHIBITORY SITE.
RX MEDLINE=81212827; PubMed=6165619;
RA Sottrop-Jensen L., Loenblad P.B., Stepanik T.M., Petersen T.E.,
RA Magnusson S., Joernvall H.;
RA "Primary structure of the 'bait' region for proteinases in alpha 2-
```



```
Dh 217 H----PFTVEEFLPKEFVQVTPKIIITILEEENMVSGLTGKPYBGHTVSICRKY 273
Qy 240 AAEIMGTDSDLMGTYQOSHMOJFNCK-NSRPRVHDYOPEITLPQVSDLPMDGKVR 298
Db 274 DASCHGEDS-----QAFCEKFSQOLNSHGCFYQOVKKVFPOLKREKEMKLTAEQ 325
Qy 299 MIGGAVHNSNG-ESAKLSRSMNR-----AYLMAGMEKNLVMRPRMGR1-FKGGSG 349
Db 326 IOEGGYVELTGRQSSSEITRTIKLSFVKYDSHRCGIPF-----FQVRLVQKGG 376
Qy 350 SOPDNDPIL-----DYG-----YGDVRFYOLENKSNSIGT-----VRYNPRS 389
Db 377 -VPIPKVIFIRGNEANYSNATDEHGLVQFSI---NTTNVNGTSLVRYNYNDRS 429

RESULT 15
INVO_MOUSE STANDARD; PRT; 467 AA.
ID INVO_MOUSE
AC P48997;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE INVOLUCRIN.
GN IYL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH SWISS;
RX MEDLINE=94104476; Pubmed=8277848;
RA Djian P., Phillips M., Easley K., Huang E., Simon M., Rice R.H.,
RA Green H.;
RA "The involucrin genes of the mouse and the rat: study of their shared
RT repeats.";
RL Mol. Biol. Evol. 10:1136-1149(1993).
CC -I- FUNCTION: INVOLUCRIN IS A KERATINOCYTE PROTEIN THAT FIRST APPEARS
CC IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO
CC MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE
CC FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.
CC -I- TISSUE SPECIFICITY: PRESENT IN KERATINOCYTES OF EPIDERMIS AND
CC OTHER STRATIFIED SQUAMOUS EPITHELIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L28819; AAA39330.1; -.
CC DR MGD; MGI:96626; IYL.
CC DR InterPro: IPR002360; Involucrin.
CC DR PROSITE: PS00795; INVOLUCRIN; 1.
CC KM Keratinocyte: Repeat
CC SO SEQUENCE 467 AA; 54919 MW; 603E1E51B435737D CRC64;

Query Match 3.8%; Score 90.5; DB 1; Length 467;
Best local Similarity 19.8%; Pred. No. 7.1;
Matches 62; Conservative 44; Mismatches 128; Indels 79; Gaps 14;

Qy 134 PUVVY-----OETDPIFLMG-----NEKGMLTKKDAKOLEYAKOFTPLSLSFDDR 182
Db 175 POELHLRQHOEKLDLPDLHLGQOQKTPREQKLIPGEKQDELHLGQRHQEPQDELHLGOK 234
Qy 183 NTPLMSSRPNNPMVULPIFMHGKRNBSRPNTPSHAROPTRENFAPRLKFOYSYKVAKE 242
Db 235 Q----KQKLEHPELQKQKQKQKPS-BPELPKQKQKQKQKQKQKQKQKQKQKQKQKQ 285
Qy 243 DLWGTDSDLMGFYQOSHMOJFNCKNSRPRVHDYOPEITLPQVSDLPMDGKVRMIGM 302
```

```
Dh 286 -----PELQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 333
Qy 303 GAVHNSNGESAKLSRSMNRAYLMAGMEKNLVMRPRMGR1FKGGSGSOPDNDPILDY 362
Db 334 GKQKHQECQEPQL-----LEEKQKHQKPEPELHLGKQKQKQKQKQKQKQKQKQKQKQ 377
Qy 363 GYGDVRFYOLENKSNSIGTYRYNPRSGKALQLDYVPLKKGISGYRQITQGYGQSL-- 420
Db 378 -----LEEKQKLG-----BPELHLGKQKQKQKQKQKQKQKQKQKQKQKQKQ 416
Qy 421 -----IDYNH 425
Db 417 EKASRQQLDYSH 429
```

Search completed: November 30, 2001, 14:27:03
Job time: 550 sec

11 11 11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:32 Search time 44.45 Seconds
(without alignments)
757,461 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSILPCFALIA.....YNHEATSPGVGLMDMMGL 442

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: PIR68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2 E81195	phospholipase A1,
2	815	34.5	409	2 H81831	probable phospholip
3	373	15.8	286	2 B36971	outer membrane pho
4	366	15.5	329	2 D81279	phospholipase A1 (
5	358	15.2	289	2 A36971	outer membrane pho
6	357	15.1	289	1 PSCA1	phospholipase A1 (
7	357	15.1	289	1 E86069	outer membrane pho
8	321.5	13.6	289	2 C36971	outer membrane pho
9	246.5	10.4	355	2 C64582	phospholipase A1 p
10	240.5	10.2	355	2 H71930	probable phospholi
11	109.5	4.6	824	3 JC7503	cellulase (EC 3.2.
12	109	4.6	800	2 A29603	cellulase (EC 3.2.
13	109	4.6	822	2 JT0611	cellulase (EC 3.2.
14	106.5	4.5	602	1 TVRTRR	protein kinase (EC
15	101.5	4.3	783	2 JC5467	cellulase (EC 3.2.
16	98.5	4.2	1658	2 T42642	phosphoinositide 3
17	97.5	4.1	797	2 T46737	x-pro dipeptidyl-p
18	95.5	4.0	389	2 S76490	hypothetical prote
19	94.5	4.0	387	2 A35136	cellulase (EC 3.2.
20	94.5	4.0	601	2 T26062	hypothetical prote
21	94.5	4.0	655	2 D26061	hypothetical prote
22	94.5	4.0	1641	2 D82704	conserved hypothet
23	94	4.0	696	2 S55694	protein kinase (EC
24	94	4.0	765	2 T35719	chitinase - Strept
25	93.5	4.0	584	2 T40777	ferric reductase t
26	93.5	4.0	901	2 T20122	hypothetical prote
27	93	3.9	633	2 C81956	thiamin biosynthes
28	93	3.9	791	2 H96839	hypothetical prote
29	93	3.9	4273	2 C69679	polyketide synthas

30	93	3.9	4307	2 T20721	hypothetical prote
31	92.5	3.9	857	1 A41369	S-receptor kinase
32	92.5	3.9	1807	2 T30940	vitellogenin - Pim
33	92	3.9	324	2 B69521	hypothetical prote
34	92	3.9	888	2 S50801	AMP deaminase homo
35	92	3.9	2971	2 T08026	hypothetical prote
36	92	3.9	5005	2 F82884	hypothetical prote
37	91.5	3.9	766	2 JC7355	peroxisome prolife
38	91	3.9	367	2 T24058	hypothetical prote
39	91	3.9	1078	2 T18352	protein P120 - Myc
40	91	3.9	1474	1 MAHU	alpha-2-macroglobu
41	90.5	3.8	386	2 A81328	probable periplasm
42	90.5	3.8	467	1 A49377	involucrin - mouse
43	90	3.8	841	2 JC5894	killer cell inhibi
44	90	3.8	1788	2 T29043	hypothetical prote
45	89.5	3.8	477	2 S23257	alpha-amyrase (EC

ALIGNMENTS

Query Match	Score	815:	DB 2:	Length	382:
Best Local Similarity	44.7%	Pred. No. 1.4e-59:			
Matches 163:	Conservative 67:	Mismatches 111:	Indels 24:	Gaps 8:	
QY 87	INCSALNODIMRLACQDTLVHGETPAVY-----KTKRSIRLDETIMGTI-KGKQVYVQE 140				
DB 33	LOCAALTDNVTYRLACQDTLVHGETPAVY-----KTKRSIRLDETIMGTI-KGKQVYVQE 92				
QY 141	TTDPFIFLWNGEKMTLTKKDAQLEKAAQETPLSFDLIDNN-TPLMSSRPHPMYLP 199				
DB 93	GGDL-----PADSAGETADITPTLSLMDLDKNDRLGLGVREHHPMYLP 139				
QY 200	IFMKGKPNRSPNTPSSHEAR-OFTPNRFPAPLKQVSYKVAADLWGTDSDFGTQQ 258				
DB 140	LMVNNSPNYVAGSPTRGTVOEKFRATKTLQVSEKSLIADLPEKTRADLFGYTOR 199				
QY 259	SHMOIFN-GKNSRPFVHVDQPEIFLQPVYSDI-PMGKVMIMGMGVHNSGSAKLSR 317				
DB 200	SDMOIYNQGRKSAFPRNTDYKPEIFLQPVYADLPFGKRLMLAGFVHNSGSRPESR 259				
QY 318	SWNRAYLMAGMEKMLNLTVMRPRIMGRIFKESGSDPDNDILDYGYGVDFVFLYLENS 377				
DB 260	SWNRITVYAMAGMEKMLNLTVMRPRIMGRIFKESGSDPDNDILDYGYGVDFVFLYLENS 317				
QY 378	NISGTVRNPNRSGKALQDLYVPLGKIGISGFQIFQGYGSLDYNHEATSPGVGLMLN 437				
DB 318	NVYSVLRNPNRSGKALQDLYVPLGKIGISGFQIFQGYGSLDYNHEATSPGVGLMLN 377				
QY 438	DMMGL 442				

A:Cross-references: GB:AE005174; NID:g12218695; PIDN:AAG59017.1; GSPDB:GN00145; UMGF:Z55
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: *pldA*
C:Superfamily: bacterial phospholipase A1

Query Match	15.1%;	Score 357;	DB 2;	Length 289;
Best Local Similarity	39.3%;	Pred. No. 6.6e-22;		
Matches	84;	Conservative 29;	Mismatches 91;	Indels 10; Gaps 4.

OY 226 RAPELKEFVSVKVAANAEDLMGCTSDMLWGYTQOOSHWOLFENKSNRPFPVHDYOEPILFTQ 285
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 83 RKDEYKFQLSLAFPLMRIGITLGPNSVLGASTYTQKSMWGLSNSSESSPFRFETNYEBOQLFGF 142

OY 286 PV-YSDLPMDCRVRMIGCAVHSHNGESAKLSRSKNRAYLAGMEKMKLVMPRIWGRIE 344
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 143 ATDYRFAGM--TLRDVEKGYNHDSNGRSDPTSRSNRNRLYTLRLAENGMLLEVKKPMWYV- 199

OY 345 KEGSGSQDDNDPDLIDYYGYGDVREFLYOLENKSNISSGVIRKNRPNPSGKALOLDIYVPFGK 404
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 200 -----GNTDNDPDLITTKYGYQLRIGYHL-GDAYVLSAKGOYNNMTGYGGAEELGLSPYTK 253

OY 405 GISGFOLIFOGYGSLIDYHNHEATSFVGGLMD 438
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 254 HVRLTYOVISGTGESLIDYNNOTRFGVGYMLND 287

RESULT 8
C36971
outer membrane phospholipase A (EC 3.1.1.-) precursor - Proteus vulgaris
C:Species: Proteus vulgaris
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: C36971; S40J30
R:Brok, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial plda genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: C36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <BRO>
A:Cross-references: EMBL:X76902; NID:9436889; PTDN:CA54224.1; PTD:9436890
A:Note: authors translated the codon GAG for residue 74 as Gly and CGA for residue 115 a
C:Genetics:
A:Gene: plda
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match	13.6%	Score 321.5;	DB 2;	Length 289;
Best Local Similarity	33.3%	Pred. No. 5.7e-19;		
Matches 87; Conservative	42;	Mismatches 115;	Indels 17;	Gaps 7;

QY 179 LDRNNTPMSSRRPNNPVLP LIFMHGKPNRSPPNPSPHEARQFTPNEEFAPLAKQOVSKY 238
 QY 43 LQEHDPN-FTLYPESNNLTYYT-----TSDLNKKALESYMWSDNA-NKDEYKQQLSLAF 95
 QY 239 KAEDDKCTSDDLFMFGTQQSHMOIFENGKNSRPRVHVDYOPEITLQPV-YSLDLPMDGKV 297
 Db 96 PLMGILIDNLSLILASAYQSRMQLSTNGESAPAPRETTYEFQLGATDYSVDW--TL 153
 QY 298 RMIGMGAVHHNGESAKLSRSRNNRAYLLACGEMENLTPMPIMWRIFKESGSGSPDDNP 357
 Db 154 RDAEFGYVHNGSNGSDPFSRSNNRLYSRLMAQNGNMLVEYKPMVIYI-----GQTSDDKN 207
 QY 358 ILDYYGVDVFLQLENNKSNISGTVRYNPRSGALQDLVYDPLGKIGISGYFOIFQGY 417
 Db 208 ITRKMGYQLQIGIQL-GEAVLSAKGQYNNMTGYGGAGELGVSYITRKHVREYTVYSGYG 266
 QY 418 QSLIDYNNHATSPGVGLMLND 438
 Db 267 ESLIDYDENQTRVGMGVMLND 287

RESULT 9
C64582
phospholipase A1 precursor - Helicobacter pylori (strain 26695)

C:Date: 09-Aug1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: C64582
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee-
son, J.D.; Kelley, J.M.; Cotton, M.D.; Waidman, J.M.; Fujii, C.; Bowman, C.; Watthey, J.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpi, P.D.; Smith, H.O.; Fraser,
A.:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A64520; MUID:97394467
A:Accession: C64582
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-355 <TOM>
A:Cross-references: GB:AE000564; GB:AE000511; NID:92313602; PIDN:AD07564.1; PID:923213

Query Match	10.4%;	Score 246.5;	DB 2;	Length 355;
Best Local Similarity	25.1%;	Pred. No. 1.2e-12;		
Matches 82;	Conservative 41;	Mismatches 117;	Indels 87;	Gaps 11

```

OY 157 KKDKAQLEUYAAKOFPEPLSEDDLRNNRNTPLSSRPNNMYULPLFEMCKKPPRSPTPSHE 216
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 69 KKYLNMDUOLSTFLPRVHSH-----TRPGVYHNPNINP----- 102

OY 217 AKQETPRNEFRAPELKFOYUSVYUKAAEDLMTGSDLMFEGYTOQSHKQITNGKNSRPFRRYND 276
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 103 ---VORNE-----KFOISFRPVFRRLHLMWKGTLYLTLTDTDFOLYINDQSOAPMRMIN 154

OY 277 YQPEIFELPOFVYSDLRPMGVK---RMIGGAVHNHNSG-ESAKLSRSMRAVLMAMEKN 332
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 155 FEMELIUYPR-NEKRFEGKIKGNSEITWIGMHISNGVGAOCYQPFK-----EGNPNQ 209

OY 333 LTVMPRI-----WGRIFKEGSGOP-----DDNPD 357
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 210 FPGQPVYIKDYNGCKDVRWGGCRSVSAGORPVFRILWMEKGLKIMVAYMPVYVDQSNPN 269

OY 358 ILDYVGDVAFELY-----QLENKSNISGTVRYNPRSGALQOLDVYPLGKISGYF 410
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 270 LIDYMGVGNALIDYRGRGHHFELDLYDIFLQYWRD--RMHGARLRYGTYVINPFVGIYA 327

OY 411 QIFQGYGSLIDYVNHHEATSPGVGLMLN 437
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 328 QWFNGYGDGLYEYDVFSNRIGVGIRLN 354
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

```

RESULT 10
H71930
Probable phospholipase a1 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: H71930
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doly, P.C.; Smith, D.
; Ives, C.; Gibson, R.; Metherg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.E.
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
A:Reference number: A71800; MUID:99120557
A:Accession: H71930
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-355 <ARN>
A:Cross-references: GB:AE001479; GB:AE001439; NID:g4154979; PIDN:AD06029.1; PID:g415
A:Experimental source: strain J99
C:Genetics;
A:Gene: plda

QY 313 -----AKLSRSMNRAVYLMAGMEKNTVMPRTWGR 342
 Db 424 VEDVYIENEGALKLISGLDASNDVSEGNWYMANRLSADGWMK 465

RESULT 13

JT0611
 cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
 N:Alternate names: endo-1,4-beta-glucanase
 C:Species: Bacillus sp.
 C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
 C:Accession: JT0611
 R:Sunltomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
 Biosci. Biotechnol. Biochem. 56, 872-877, 1992
 A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
 A:Reference number: JT0611; MUID:92305459
 A:Accession: JT0611
 A:Molecule type: DNA
 A:Residues: 1-822 <SUN>
 A:Cross-references: GB:M04963; NID:9289264; PIDN:AAA73189.1; PID:9289266
 C:Function:
 A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel-
 A:Pathway: cellulose degradation
 C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylianse A amino-terminal ref
 C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
 F:353-726/Domain: Thermotoga xylianse A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 109; DB 2; Length 822;
 Best Local Similarity 19.7%; Pred. No. 1;

Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;

QY 30 PVAEVDVRSKND-----LGQDNELLIGVQATQASATDTANPL-DEHEPELY-----TT 78
 Db 25 PALAAEGNTREDNFKHLGNDNRKPSRSEAGALQLEVDGQGTIVDQHEKIQLRGMSH 84
 QY 79 ALENTMLINCASL-----NQDIMRLACY-DTLVHGFTPAVITKRSIRIDETI----- 126
 Db 85 GLQMPRELLINDNAYKALANDWESNMIRLAMYGENGYVSNPELISRYIKGIDLAENDM 144
 QY 127 -----WQF-----IKGPPQVVOETDPP-----IFLMGN 150
 Db 145 YVIVDMVHAPRDPDPVYAGAEDEFROIALYPRNNPHIYLANPSSNNNGAGIRPN 204
 QY 151 EKGMLTKKDAKOLEYAAKOPTLSLFDLRN-----NTPLMSSRP-----HNPMY 196
 Db 205 EEGMNAVK-----EYADPTVEMLRDSGNADNIIIVGSPNMGQRPDLADNPIIDHHTWY 259
 QY 197 VLPPIF--MHGKPNRS--PMTPSHEARQFTPNFRAPELKFQVSVKVKAAEDLMGT----- 247
 Db 260 TVHETGSHIASTESYPPETPNSEKGNWMSNTRYA-----LENGVAVFATE--WGTSQLANG 313
 QY 248 -----DSDLWFGYTOQ-----SHMOIFNGKNS-----RPPRVH----- 275
 Db 314 DCGPYFDEADVIEFLNENISMAMNSLTN--KNEVSGAFTPELKSNAATSLDPPGQYW 372
 QY 276 -----DYOPEITLQPVYSDLPW--DGKVMNMGCAVHSHNGES- 312
 Db 373 VPEELSLGSEYVRARIKGVNPER--IDRTKYTKVLMDFNDCTKQGFV-----NGDSP 423
 QY 313 -----AKLSRSMNRAVYLMAGMEKNTVMPRTWGR 342
 Db 424 VEDVYIENEGALKLISGLDASNDVSEGNWYMANRLSADGWMK 465

RESULT 14

TVRTRR
 protein kinase (EC 2.7.1.37) raf - rat
 N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-8
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
 C:Accession: B26126

R: Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
 Mol. Cell. Biol. 7, 1226-1232, 1987
 A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused prote
 A:Reference number: A26126; MUID:87172791
 A:Accession: B26126
 A:Molecule type: mRNA
 A:Residues: 1-602 <ISH>
 A:Cross-references: GB:M15428; NID:9206546; PIDN:AAA42002.1; PID:9206547
 C:Genetics:
 A:Gene: raf
 C:Superfamily: rat protein kinase raf; protein kinase homology
 C:Keywords: Atp; autophosphorylation; phosphoprotein; phosphotransferase; serine/thre
 F:301-567/Domain: protein kinase homology <KIN>
 F:309-317/Region: protein kinase Atp-binding motif
 F:329/Active site: Lys #status predicted
 F:453/Binding site: phosphate (ser) (covalent) (by autophosphorylation) #status predi

Query Match 4.5%; Score 106.5; DB 1; Length 602;
 Best Local Similarity 19.0%; Pred. No. 1;

Matches 92; Conservative 74; Mismatches 172; Indels 145; Gaps 24;

QY 23 QOAVPNPVAFVDEVRSKNDLGQDNELLIGVQATQASATDTANPLDEHEPELYTTALEN 82
 Db 81 ELEKLNQYKVTLEKNKELETQADNRN--GISOFTRAEE-----LEAKRLIRT--N 131
 QY 83 KTMILNCASLQNDIMRLACYDTLVHGETPAVITKRSIRIDETIQ--TIKGPQVVOYE 140
 Db 132 ERLSQEVEYLTEDVRKL--NEKLRESNT--TKGEIQTLKDELQASDVYVKYREKRLEQE 186
 QY 141 -----TTDPIFLMGNEKG-----MLTKKDAKOLEYAAKOPTPLS 174
 Db 187 KELLNQNSWLTETKTDDELALGREGNEILEKCTLENKKEDAIRSHSESASPSA 246
 QY 175 ISFDDLRNNTPLMSSRPNPNAYVLPIFMHGKPNRSPTPSHAROPTPEFAPAPELKFOY 234
 Db 247 LS-SSPNMSPGQWQPKRP-----VPAQREARPGSOTGQKNKIRPGQDSSYWEI 298
 QY 235 SVKVKAAEDLQCTSDSLMEFGYQOSH-----QIFNGKNSPPRYVDYQPEI----- 281
 Db 299 EASEVMLSTRIGSGS--FGYIVKGMHGDVAVKILKVDPTPEQLQAFRNEVAVLRKTR 355
 QY 282 -----FLTO-----PYVSDL--PYMDGKVRML-----GMGAVNH 307
 Db 356 HVNILLPMGYMTKMDLAITYQWCEGSSSLKHLHVOETKFOMLQIDIAQOTAGMDYLHA 415
 QY 308 SNGESAKLSRSMNRAVYLMAGMEK-----NLTVMPRTWGRIFKESGSGQ----- 351
 Db 416 KNIIRHDKMS--NITFLHGLTVKIGDGLATVYKSRM-----SGSQQVEQPTGSVLM 466
 QY 352 -----PDNDPDL--DYGGYGVRELYQLENKSNISGTVRYNPRSGKGLQLDYVPL 402
 Db 467 APEVIRMODNPNFSPQSDVYSYIV--LYEL-----MTGELPYSHINNDDI-----IFWV 515
 QY 403 GKG 405
 Db 516 GRG 518

RESULT 15

JC5467
 cellulase (EC 3.2.1.4) - Bacillus sp.
 N:Alternate names: Endo-1,4-beta-glucanase
 C:Species: Bacillus sp.
 C:Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 10-Dec-1999
 C:Accession: JC5467
 R: Miyatake, M.; Imada, K.
 Biosci. Biotechnol. Biochem. 51, 362-364, 1997
 A:Title: A gene encoding endo-1,4-beta-glucanase from Bacillus sp. 22-28.
 A:Reference number: JC5467; MUID:97212038
 A:Accession: JC5467
 A:Molecule type: DNA
 A:Residues: 1-783 <MIY>

A:Cross-references: DDBJ:D85236

A:Experimental source: strain 22-28

A:Note: neither the complete nucleic acid sequence nor the complete translation are shown

C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal rep

C:Keywords: glycosidase; hydrolase

F:587-728/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.38; Score 101.5; DB 2; Length 783;

Best Local Similarity 19.7%; Pred. No. 3.9; Mismatches 148; Indels 169; Gaps 26;

Matches 91; Conservative 54; Mismatches 148; Indels 169; Gaps 26;

```
QY 30 PVAFVDEVRKND-----LGODNELLGVSANQSASTDANPL-DEHEPELY-----TT 78
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 25 PTALAEAGNTREDNFHLLNENYKRPSEAGALQKVEYDGMITLVQHGKIQLRGMSTH 84
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 79 ALENKTMILNCSAL-----NODIMRLACYDTLVHGE-----TPAVIKTK-----RSI 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 85 GLQWFPEILNDNWKALSDNWDNMIRLAMY---VGENGYATNPBELIKQVIDGIELAI 140
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 RLDETI---WQT-----IKGRQVYVYQETDP-----IF 146
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 141 ENDMYVIVDMHVAHPGDPDPVYAGAEDEFRDIAALYPNMNPHTIYELANEPSSNNNGAG 200
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 147 LMGNEKMLTKDAKOLEYAAKQFTPLSLSFDDLRN---NTPLASSRP-----H 192
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 201 IPNNEEGWKAVK-----EYADPIYEMLRDSGNADNIIIVGSPNMSQRPDLADNPINDH 255
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 193 NPMYVLPDI--MHGKPNRS--PMTPSHEARQFTPNNEFRAPELKFQVSVKYKAAEDLWGT- 247
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 256 HTMYTVHFYTGSHASTESYPPETPNSEKGNVMSNTRYA---LENGVAVFATE--WGTS 309
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 248 -----DSDLWEGYTOQ-----SHWQIFNGKNS-----RPRV----- 274
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 310 QANGDGGPYDEADWIEFLNENNISWANNSLTN-KNEVSGAFTPELIGKSNATSLDGP 368
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 275 -----HDYQPEIFLTQPVYSDLPW---DGKVRMIGMGA----- 304
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 369 DQWVAPEELSLSGEYVARIKAKYEP---IDRTRYTKVLMDFNDDGTRKQGFVNSDSPNK 425
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 305 ----VHHSNGESAKLSRSNRRAYLMAGMEWKNLTVMPRINGR 342
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 426 EAIEVENENG-TLRIISGLNVSNDLSGDNFWANFRLSANGWGK 466
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 30, 2001, 14:18:34
Job time: 256 sec

This Page Blank (uspto)

NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-13

Query Match 4.2%; Score 99.5; DB 2; Length 1658;
Best Local Similarity 18.6%; Pred. No. 0.65;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

QY 1 MKVSL-----TLTLSPCFALLAIQQAQAVNPVAVD-----EYRSKNDLGQD 46
DB 397 VKVSIIEGQLPYTFCDVSVTEIIMQALCWVHDDLQVDGSYTLKCGGEEVLQN 456
QY 47 NELLIGVQSATQSTPTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDPLV 106
DB 457 NHCLGSHHIONCKRMT-----EIKQLTLTSLAMCONLARTEDD--- 497
QY 107 HGETPAVVKTRIRLDETITQIKGRPOVYVOETDPIF----- 146
DB 498 --EAP-----VDLKNYLYQIEKPYKEVMIRHVEELDSYHYVELALQTEHQRAV 547
QY 147 -----LMGNEKMLTKKDAKOLEYAA-----KQFTPLSLSFDDRNTPLMS 188
DB 548 DQVIKAVRKICSAIDGVEVPSVTEA-VKLLKRAVNLPRNKSADVTSLSGSDTRKNSKGS 606
QY 189 SRPNPMYVLP-----IFMHGKPNR-----SPNTPSHARQFPPNEFRABELKQV 234
DB 607 LNPNPQVQSMHDLTTALYYDLRLHANSRCSGCPGRSRIKAWATE-----QL 658
QY 235 SVKKAEDL--WGTDSDLMFGYTOQSHMOIFNGKN--SRPF--RVHNDYQPEIFLQ-- 285
DB 659 QFTVYAAHGISSNMVSNVEKYLLCSLSH-----NGKDLFRQISKVGTYNFYLKMD 714
QY 286 -----PV-YSDLPMDGKVRMIGMGAHVHNSGESAKLSRSMN-----RAYL 324
DB 715 ELITFIQISQLPLESVLHTLFGVLNQSSGSSPDNSKQKRGPEALGVSLTLDFKFRFL 774
QY 325 MAGMEMKNTLVMPRIWGRIFKEGSGSQDDNPDLIDYGYGDVDFLYOLEKNSISGTVR 384
DB 775 TCG-----TKLTYLW-----TSSHTNSIPGAIIPKKSVMERIVLQVDPPSPAFDIIY 821
QY 385 YNPRSGKALQLDVYVPLGKISGYFOIFQGYGSLIDYNEHATSGV 432
DB 822 TSPQIDRNIIQODKLETLESIDIKG-----KLDDIHRDSSFGL 859

RESULT 4
US-09-170-996-13
Sequence 13, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moitz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.2%; Score 99.5; DB 4; Length 1658;
Best Local Similarity 18.6%; Pred. No. 0.65;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

QY 1 MKVSL-----TLTLSPCFALLAIQQAQAVNPVAVD-----EYRSKNDLGQD 46
DB 397 VKVSIIEGQLPYTFCDVSVTEIIMQALCWVHDDLQVDGSYTLKCGGEEVLQN 456
QY 47 NELLIGVQSATQSTPTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDPLV 106
DB 457 NHCLGSHHIONCKRMT-----EIKQLTLTSLAMCONLARTEDD--- 497
QY 107 HGETPAVVKTRIRLDETITQIKGRPOVYVOETDPIF----- 146
DB 498 --EAP-----VDLKNYLYQIEKPYKEVMIRHVEELDSYHYVELALQTEHQRAV 547
QY 147 -----LMGNEKMLTKKDAKOLEYAA-----KQFTPLSLSFDDRNTPLMS 188
DB 548 DQVIKAVRKICSAIDGVEVPSVTEA-VKLLKRAVNLPRNKSADVTSLSGSDTRKNSKGS 606
QY 189 SRPNPMYVLP-----IFMHGKPNR-----SPNTPSHARQFPPNEFRABELKQV 234
DB 607 LNPNPQVQSMHDLTTALYYDLRLHANSRCSGCPGRSRIKAWATE-----QL 658
QY 235 SVKKAEDL--WGTDSDLMFGYTOQSHMOIFNGKN--SRPF--RVHNDYQPEIFLQ-- 285
DB 659 QFTVYAAHGISSNMVSNVEKYLLCSLSH-----NGKDLFRQISKVGTYNFYLKMD 714
QY 286 -----PV-YSDLPMDGKVRMIGMGAHVHNSGESAKLSRSMN-----RAYL 324
DB 715 ELITFIQISQLPLESVLHTLFGVLNQSSGSSPDNSKQKRGPEALGVSLTLDFKFRFL 774
QY 325 MAGMEMKNTLVMPRIWGRIFKEGSGSQDDNPDLIDYGYGDVDFLYOLEKNSISGTVR 384
DB 775 TCG-----TKLTYLW-----TSSHTNSIPGAIIPKKSVMERIVLQVDPPSPAFDIIY 821
QY 385 YNPRSGKALQLDVYVPLGKISGYFOIFQGYGSLIDYNEHATSGV 432
DB 822 TSPQIDRNIIQODKLETLESIDIKG-----KLDDIHRDSSFGL 859

RESULT 5
US-07-717-331F-2

```

; Sequence 2, Application US/0717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717.331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-717-331F-2

Query Match          3.8%; Score 89.5; DB 1; Length 857;
Best Local Similarity 21.4%; Pred. No. 2.5;
Matches 76; Conservative 50; Mismatches 110; Indels 119; Gaps 23;

QY 105 LVHGERPAVIAKTRSRIRLEETIMQTKGRPOVYQETTPIDPLMG---NEK---GMLTK 157
DB 24 LHPALSIITNT---LSTESLTSSNKTLY---SPGSIIEVGFEFRNRSRWYLGMMYK 75
QY 158 K-DAKOLEYAAKQFTPLSLSPD-----LDRNNTPLMSSRPHNPMYVLPFMHGK 205
DB 76 KYSDRIYVAVARNRDNPLSLAIGTLKISGNNILVLDHSNKPVMWTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEARQFT-----PNEFRAPDELKFOVSVKAAEDIM 245
DB 126 -RSRPAVIELLANGNEVMDSSNNDSSEYLMOSFDYPTDTLLPEMKLGYNLKT----- 177
QY 246 GTDSDMFEGYTOOSHQIENGKNSRPERVHDYQ-----PEIFLQ-----PYVSDLPMDCK 296
DB 178 GLNRFL-----TSNRSDDPSSGNF---SYKLEQTQSILPEFLSKENRPMHRSQPMWNG- 226
QY 297 VEMIGGAVH-----HSNGESAKLSRSMNRAY-----LMAGMEKNTLVMP--RI 339
DB 227 IIFSGIPEDQKLSYMYNFIENNEEVAAYFRMTNNSFGRLLSLISGIRQLRWPSIRI 286
QY 340 WKRIRKESGSGPDD---NPDILDYGYGDVRFELYQLENKSNISGTVR--YMPRS 389
DB 287 NMRFSSPVDROCDIYIMGCP-----YAYCDV-----NTSPVCNCIOGFNPRN 329

```

```

; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Misset, Onno
; APPLICANT: Van der Laan, Jan M.
; APPLICANT: Lenting, Herman B.M.
; TITLE OF INVENTION: Mutated beta-lactam acylase genes
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GOMARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (Epo)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/731,157A
; FILING DATE: 19910509
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90200962
; FILING DATE: 18-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER PH.D., BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: GBRO-027/00US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 844 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Kluyvera citrophila
; STRAIN: ATCC 21285
; US-07-731-157A-6

Query Match          3.7%; Score 87.5; DB 1; Length 844;
Best Local Similarity 21.2%; Pred. No. 3.9;
Matches 92; Conservative 46; Mismatches 152; Indels 143; Gaps 25;

QY 31 VAFVDEVRSK--NDLQGD--NEL--LIGVOSATQASSTDTFANPLDEHEPELYTTAL----- 80
DB 182 LALLTAVKDKYNDDEGMAVFNOLKWLNPASPTTIARESSYPLKFDLQNTOTALLVPR 241
QY 81 -ENKTMILINCSALNDIMLACYDILVHETPAVIAKTRKISILDETIWQTIKGPQVYQ 139
DB 242 YQAPAMLDLRAKGTGALLAY-----TAIKNRRTIAA-----OPANGANGLACY 286
QY 140 ETTDPIFLMGNEKGM/LTKDAKOLEYAAKQFTPLSLSPDLDRNNTPLMSSRPHNPMYVLP 199
DB 287 PTTSMNVAVIGKKA---QDAKAIWNGPFG-----W---YAPAYTYG 323
QY 200 IFMHGKP-NRSPNTSHEARQFTPNEFRAPDELKFOVSVKAAEDIMGT-----DSDIM 252
DB 324 IGIHAGYDVGTNTP-----FAYPGLVFGHNGTIS-----WGSTAGFGDQVDVIF 367
QY 253 F-----GYTQOSHWOIFENGKNSRPERVHDYQPEIFL-----TQPVYS 289
DB 368 AEKLSAEKPGYTOHNGEWKMLSRKETTIAVKDQPEFTYVMTLDSGNVIKTTPTROTATA 427
QY 290 DL-PMDGKVRMIGAVHHSNGESAKLSRSMNRAYLMAGMEKNTLV-----VMPRIW-- 340
DB 428 KARAWAGK-----EVAASL-LAW--THOMKAKKNPENTQQAQKALTIMWY 470

```


RESULT 12
 US-08-483-322-2
 : Sequence 2, Application US/08483322
 : Patent No. 5760178
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: HERRLICH, Peter
 : APPLICANT: PONTA, Helmut
 : APPLICANT: GUENTHER, Ursula
 : APPLICANT: MATZKU, Siegfried
 : APPLICANT: WENZL, Achim
 :
 : TITLE OF INVENTION: VARIANT CD4 SURFACE PROTEINS, DNA
 : TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS
 : AND AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
 :
 : NUMBER OF SEQUENCES: 8
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Foley & Lardner
 : STREET: 3000 K Street, N.W., Suite 500
 : CITY: Washington, D.C.
 :
 : COUNTRY: USA
 :
 : ZIP: 20007-5109
 :
 : COMPUTER READABLE FORM:
 :
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/483,322
 : FILING DATE: 07-JUN-1995
 : CLASSIFICATION: 435
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/946,497
 : FILING DATE: 09-NOV-1992
 :
 : ATTORNEY/AGENT INFORMATION:
 :

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:16:58 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 2360

Sequence: 1 MKVSLSTFLSLIPCFATLA.....YNHEATSRGVLMLNDMWGL 442

Scoring table:

BLOSUM62

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq_1101:*

- 1: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2360	100.0	442	21	BAAB034 amino acid
2	2339	99.1	442	21	BAAB034 amino acid
3	2330	98.7	442	21	BAAB034 amino acid
4	2330	98.7	442	21	BAAB034 amino acid
5	822	34.8	370	21	BAAB034 amino acid
6	815	34.5	370	21	BAAB034 amino acid
7	815	34.5	370	21	BAAB034 amino acid
8	815	34.5	370	21	BAAB034 amino acid
9	812	34.4	375	21	BAAB034 amino acid
10	246.5	10.4	355	19	AAV70628
11	239.5	10.1	356	19	AAV10960

12	156.5	6.6	253	18	AAW20760	H. pylori outer me
13	109.5	4.6	824	21	AAAB23180	Bacillus sp. KSM-S
14	109	4.6	800	8	AAV70420	Sequence encoded b
15	109	4.6	822	13	AAAB26021	Alkaline cellulase
16	107	4.5	157	18	AAW20538	H. pylori outer me
17	106	4.5	798	21	AAAB40925	Human ORFX ORF689
18	106	4.5	798	21	AAV90225	Human PGC-1 protei
19	106	4.5	798	21	AAAB3944	Peroxisome prolif
20	105.5	4.5	537	22	AAAG1000	C glutamicum prote
21	102.5	4.3	166	19	AAV70991	Amino acid sequenc
22	98.5	4.2	467	20	AAV13378	Human class II PI3
23	98.5	4.2	467	21	AAV94860	Human protein clon
24	98.5	4.2	467	22	AAAB88401	Human membrane or
25	98.5	4.2	467	22	AAAB80246	Human PRO260 prote
26	98.5	4.2	1726	18	AAAB38756	Phosphatidyl inosi
27	97	4.1	761	20	AAW99084	Non-B, non-C, non-
28	95.5	4.0	502	22	AAAG90031	C glutamicum prote
29	95.5	4.0	516	22	AAAG90031	C glutamicum prote
30	95	4.0	522	22	AAAB95616	Human protein sequ
31	94	4.0	372	16	AAAB69607	Human protein sequ
32	93.5	4.0	467	19	AAW75057	Human secreted pro
33	92	3.9	888	22	AAAG70751	S cerevisiae apopl
34	92	3.9	1484	12	AAAB11749	Human alpha-2 macr
35	91.5	3.9	857	13	AAAB29814	S receptor kinase
36	91	3.9	600	22	AAAG75092	Human colon cancer
37	91	3.9	1227	22	AAAB81501	S. epidermidis ope
38	91	3.9	1474	21	AAV97157	Human alpha-2-macr
39	91	3.9	1474	21	AAAB50673	Human alpha-2 macr
40	90	3.8	522	22	AAAB95513	T. matsutake pyran
41	90	3.8	564	21	AAAB10457	Trichoderma deriva
42	90	3.8	564	21	AAAB1952	Human protein sequ
43	90	3.8	760	22	AAAB93045	Human dephosphoryl
44	90	3.8	841	20	AAV24318	Mouse dephosphoryl
45	90	3.8	1627	16	AAAB67538	Cytadhesin Pl. My

ALIGNMENTS

RESULT 1

AAV85269 standard; Protein: 442 AA.

AC AAV85269;

XX 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #2.

XX DE

XX Moraxella catarrhalis infection; BASB034; diagnosis; straging;

KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;

KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;

KW hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

XX WO200015802-A1.

XX PD

XX 23-MAR-2000.

XX PF

XX 14-SEP-1999; 99MO-EP06781.

XX PR

XX 14-SEP-1998; 98GB-0020002.

XX (SMK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Ruelle J;

XX WPI; 2000-271440/23.

XX N-PSDB; AAA10701.

XX Novel BASB034 polynucleotides and polypeptides from Moraxella

PT catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 67; 106pp; English.
 PS
 XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderly, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 CC
 XX
 XX
 SQ Sequence 442 AA;

Query Match 100.0%; Score 2360; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 2.8e-219;
 Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTSLTPCFALITAOAOAVPNPVAFVDEVSKNDLGDNDNELLIGVQSATQSA 60
 DB 1 mkvslstltsltpcfalitaqagavpnvafvdevskndlgdndnelligvqsatqsa 60
 QY 61 STDTPANPLDEHEPELYTALENKMTLINCASALNODIMRLACYDTLVHGEPYAVITKRSI 120
 DB 61 stdtanpldehepeelytalenkmtlincsalnodimrlacydltlvhgcpaviktksi 120
 QY 61 stdtanpldehepeelytalenkmtlincsalnodimrlacydltlvhgcpaviktksi 120
 DB 61 stdtanpldehepeelytalenkmtlincsalnodimrlacydltlvhgcpaviktksi 120
 QY 121 RLDETITWQITKGPVVYQETTPITFLMGNEKMTLTKKAKOLEYAAKQFTPLSLSFDD 180
 DB 121 rldetiwtqitkgpvyvqettpitflmgnekmtltkakoleyaakqftplslsfdd 180
 QY 121 rldetiwtqitkgpvyvqettpitflmgnekmtltkakoleyaakqftplslsfdd 180
 DB 121 rldetiwtqitkgpvyvqettpitflmgnekmtltkakoleyaakqftplslsfdd 180
 QY 181 RNNTPLMSSRPNNPMTVLPITFMHGKPNRSPTNSHEARQFTNEFRAPELKQVSYKVA 240
 DB 181 rnntpmlssrpnnpmtvlpitfmhgkpnrsptnshearoftnefrapelkqvsvkva 240
 QY 241 AEDLWGTDSDLFGYITQOSHWOIFNGKNSRPRVNDYOPETFLQPVSYDLPWDGKVRMT 300
 DB 241 aedlwgtdsdlfgyitqoshoifngknsrprvndyopetflqpvsydlpwwdgkvrmt 300
 QY 241 aedlwgtdsdlfgyitqoshoifngknsrprvndyopetflqpvsydlpwwdgkvrmt 300
 DB 241 aedlwgtdsdlfgyitqoshoifngknsrprvndyopetflqpvsydlpwwdgkvrmt 300
 QY 301 GNGAVHNSGSESAKLSRSNRAVYLMAGMEKMLTVMPTRMGRIFRSGSQSDPDNDITD 360
 DB 301 gngavhnsgeesaklsrsnraylmagmekmltvmptrmgrifrgsgsqsdpdnditd 360
 QY 361 YVYGVDVRELVOLENKSNISGTVRYNPRSGKALQLDVYVPLGKGISGVFOIGYGSGL 420
 DB 361 yvgydvdrflvqlenknsisgtrvynprsgkalqldvyvplgkgsigvfgfygysgl 420
 QY 421 IDYNHEATSFYGLMLNDMGMCL 442
 DB 421 idynheatsfygylmndwmg1 442

RESULT 2
 ID AAY85271
 XX AAY85271 standard; Protein; 442 AA.
 XX
 XX AAY85271;
 XX

DT 29-JUN-2000 (first entry)
 DE BASB034 amino acid sequence #4.
 XX
 XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 XX Moraxella catarrhalis.
 OS
 XX
 XX WO200015802-A1.
 PN
 XX
 XX 23-MAR-2000.
 PD
 XX
 XX 14-SEP-1999; 99WO-EP06781.
 PF
 XX
 XX 14-SEP-1998; 98GB-0020002.
 PR
 XX
 XX (SMIT) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 XX
 XX Ruelle J;
 PI
 XX
 XX WPI: 2000-271440/23.
 DR
 XX
 XX N-PSDB; AAA10703.
 DR
 XX
 XX
 PS Claim 3; Page 69; 106pp; English.
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections -
 XX
 XX

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2969. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderly, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 CC
 XX
 XX
 SQ Sequence 442 AA;

Query Match 99.1%; Score 2339; DB 21; Length 442;
 Best Local Similarity 99.1%; Pred. No. 3e-217;
 Matches 438; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTSLTPCFALITAOAOAVPNPVAFVDEVSKNDLGDNDNELLIGVQSATQSA 60
 DB 1 mkvslstltsltpcfalitaqagavpnvafvdevskndlgdndnelligvqsatqsa 60
 QY 61 STDTPANPLDEHEPELYTALENKMTLINCASALNODIMRLACYDTLVHGEPYAVITKRSI 120
 DB 61 stdtanpldehepeelytalenkmtlincsalnodimrlacydltlvhgcpaviktksi 120
 QY 61 stdtanpldehepeelytalenkmtlincsalnodimrlacydltlvhgcpaviktksi 120
 DB 61 stdtanpldehepeelytalenkmtlincsalnodimrlacydltlvhgcpaviktksi 120
 QY 121 RLDETITWQITKGPVVYQETTPITFLMGNEKMTLTKKAKOLEYAAKQFTPLSLSFDD 180
 DB 121 rldetiwtqitkgpvyvqettpitflmgnekmtltkakoleyaakqftplslsfdd 180


```

OY 181 RNNTPLMSSRRPHNPMVLPFIEMHGKPNRSPNTPSHEAROPTPNEFRAPELKFOVSVKVA 240
    |||
DB 181 rmtptwssrrphnmpvylpifmngkpnrsprntpsheakqftptnefrapelkfvsvkva 240
OY 241 ABDLMGTDSDLMFGYQOQSHWOIFNGKNSRPFVHDYQPEIFLTQVYSDLPMDGKVRMT 300
    |||
DB 241 aedltwtdsdltwfygqgshwqifngknsrpfvhdypeliltqpyesdlpmdgkvrmt 300
OY 301 GNGAVHNSGESAKLRSRNNRAYLMAEMKNTVMPRIWGRIFKFGSGSQPDNDPDIID 360
    |||
DB 301 gngavhnsgeaklsrsnraylmagmewknltpmriwgrifkfgsgsqpddndpdiid 360
OY 421 YYGVDVRFLEYOLENKSNTSGTVRNPBSGKALQLDVYVPLGKISGYFOIFOGYGOSL 420
    |||
DB 421 yygvgdvrflyqlemnksnsgtvrynprsgkalqldvvyplgkisyfqlfygysql 420
OY 421 IDYNHEATSEFGVGLMDNMWGL 442
    |||
DB 421 idynheatsfgyglmdnmwgl 442

RESULT 3
AAV85268 standard; Protein; 442 AA.
ID AAV85268;
AC AAV85268;
XX
XX 29-JUN-2000 (first entry)
DT
XX
XX BASB034 amino acid sequence #1.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
XX OS Moraxella catarrhalis.
XX
XX PN WO200015802-A1.
XX
XX PD 23-MAR-2000.
XX
XX PF 14-SEP-1999; 99WO-EP06781.
XX
XX PR 14-SEP-1998; 98GB-0020002.
XX
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX PI Ruelle J;
XX
XX DR WPT: 2000-271440/23.
XX
XX DR N-PSDB; AAA10700.
XX
XX PT Novel BASB034 polynucleotides and polypeptides from Moraxella
XX catarrhalis used to prepare vaccines against bacterial infections
XX
XX PS Claim 3; Fig 2; 106pp; English.
XX
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
XX strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
XX from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
XX BASB034 polynucleotides and polypeptides may be employed as research
XX reagents and material for the discovery of treatments and diagnostics for
XX diseases, particularly human diseases. They are particularly used to
XX diagnose and treat M. catarrhalis infections. They can be used for
XX diagnosis of disease, staging of disease, or determining response of an
XX infectious organism to drugs. The polynucleotides may be used as a source
XX for hybridization probes, and for screening of genetic mutations,
XX serotype, organism or strain identification. Identification of mutations
XX in BASB034 sequences, and as components of arrays which are useful for
XX diagnostic and prognostic purposes. The polypeptides can be used to
XX produce antibodies. The polypeptides can also be used in vaccine

```

```

CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderly, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
XX SQ Sequence 442 AA:
OY
XX
XX Query Match 98.7%; Score 2330; DB 21; Length 442;
XX Best Local Similarity 98.6%; Pred. No. 2,2e-216;
XX Matches 436; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 1 MKVSLSTLTSLIPCFALLAIQQAQAVPNPVPVDEVRSKNDLGODNELIGVQSAQSA 60
    |||
DB 1 mkvslstltsllscfallaiqgaqavnpvafdevrseindlgdnelpidvgsatgsa 60
OY 61 STDTPANPLDEHEPELYTTALENKTMLINCSALNODIMRLACYDTLVHGEPVAVIKTKRST 120
    |||
DB 61 stdtampldhepelyttalenktmlincsalngdlmrlyacydtlvhgetpaviktkrst 120
OY 121 RLDEFTWQRIKGRQYVYQOTTPDIFLMGNEKGMILKKAKQLEFYAKQTPRISLSTFDLD 180
    |||
DB 121 rldetwqrlkgkqpyvlyqettdpdlfmgnekgmilkkakqleyaakqtprlstlstdld 180
OY 181 RNNTPLMSSRRPHNPMVLPFIEMHGKPNRSPNTPSHEAROPTPNEFRAPELKFOVSVKVA 240
    |||
DB 181 rmtptwssrrphnmpvylpifmngkpnrsprntpsheakqftptnefrapelkfvsvkva 240
OY 241 AEDLMGTDSDLMFGYQOQSHWOIFNGKNSRPFVHDYQPEIFLTQVYSDLPMDGKVRMT 300
    |||
DB 241 aedltwtdsdltwfygqgshwqifngknsrpfvhdypeliltqpyesdlpmdgkvrmt 300
OY 301 GNGAVHNSGESAKLRSRNNRAYLMAEMKNTVMPRIWGRIFKFGSGSQPDNDPDIID 360
    |||
DB 301 gngavhnsgeaklsrsnraylmagmewknltpmriwgrifkfgsgsqpddndpdiid 360
OY 361 YYGVDVRFLEYOLENKSNTSGTVRNPBSGKALQLDVYVPLGKISGYFOIFOGYGOSL 420
    |||
DB 361 yygvgdvrflyqlemnksnsgtvrynprsgkalqldvvyplgkisyfqlfygysql 420
OY 421 IDYNHEATSEFGVGLMDNMWGL 442
    |||
DB 421 idynheatsfgyglmdnmwgl 442

RESULT 4
AAV85270 standard; Protein; 442 AA.
ID AAV85270;
AC AAV85270;
XX
XX 29-JUN-2000 (first entry)
DT
XX
XX BASB034 amino acid sequence #3.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
XX OS Moraxella catarrhalis.
XX
XX PN WO200015802-A1.
XX
XX PD 23-MAR-2000.
XX
XX PF 14-SEP-1999; 99WO-EP06781.
XX

```

```

PR 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX RuelLe J:
XX WPI: 2000-271440/23.
XX N-PSDB: AAA10702.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
XX catarrhalis used to prepare vaccines against bacterial infections
XX
XX Claim 3; Page 68; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
XX strain Mc2913. The invention relates to BASB034 polypeptides from
XX M. catarrhalis strains Mc2913, Mc2908, Mc2913 and Mc2969. The BASB034
XX polynucleotides and polypeptides may be employed as research reagents and
XX material for the discovery of treatments and diagnostics for diseases,
XX particularly human diseases. They are particularly used to diagnose and
XX treat M. catarrhalis infections. They can be used for diagnosis of
XX disease, staging of disease, or determining response of an infectious
XX organism to drugs. The polynucleotides may be used as a source for
XX hybridization probes, and for screening of genetic mutations, serotype,
XX organism or strain identification, identification of mutations in BASB034
XX sequences, and as components of arrays which are useful for diagnostic
XX and prognostic purposes. The polypeptides can be used to produce
XX antibodies. The polypeptides can also be used in vaccine formulations,
XX and to identify agonists and antagonists. The polypeptides, antibodies,
XX agonists and antagonists (which are bacteriostatic) are used for the
XX treatment and prevention of diseases such as otitis media in infants and
XX children, pneumonia in elderly, sinusitis, nosocomial infections and
XX invasive diseases, and chronic otitis media with hearing loss. The
XX polypeptides, agonists and antagonists are also used for screening of
XX antibacterial drugs. The BASB034 products of the invention can be used
XX screen for new antibacterial compounds that may target resistant
XX bacteria.
XX
XX Sequence 442 AA:
SQ
Query Match 98.7%; Score 2330; DB 21; Length 442;
Best Local Similarity 98.9%; Pred. No. 2,2e-216;
Matches 437; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 MKVSLSTLTSLPCFALLAIQQAAPNPAFVDEVRSKNDLGDNDNELLVGQATOSA 60
DB 1 mkvslstltlslscfallaiaqkavnpvafvdevrsendlgqdnelpidvgsatgsa 60
OY 61 SUTDTANPIDEHEPELTYTALEKTKLINSALNODIMRLACVDTLVHGTPRAVITKRSI 120
DB 61 stctanpidehepeltytalenkmlncsalngdimrlacvdtlvhgtpaviktksi 120
OY 121 RLDEETIMOTIKKPOVVYQETTDPIFLMGNEKGMLETKKAKOLEYAAAOFTPLSFD 180
DB 121 rldetlwtikgkpvvyqettdpiflmgnekgmltkkdakleyaakqfiprlsifd 180
OY 181 RNNPFLMSSRPHPNMYVLPFEMHGKPNNSPNTPSHEAOKFTPNFRAPELKFQVSVKYA 240
DB 181 rnnpflmssrphpnmyvlpfemhgpnsnpsntpsheaoftpnfrapeklfqvsvkya 240
OY 241 AADLWGTOSDLMEFGYTOOSHMOIFNGKNSRPVRHDXQPEITLQPVYSDLPWDSKVRKI 300
DB 241 aedlwtgtsdlmefgytqshwqifngknsrptvndyqpeitlfpvydsdipwdskvrti 300
OY 301 GNGAVHSHNGESAKLSRSNRAVYLAAGMEKMLTVMPTWGRIFEGSGSQPDNDPID 360
DB 301 gngavhshngesaklsrsnraylimgemwknltvmptwgrlfegsgsqdndpid 360
OY 361 YYGVDVRFVLVYLENKSNTISGTVRYNPNRSGKALDLDVYVPLGKISGYSQITFGYGS 420
DB 361 yygvdvrfvllylenksntisgtrvrynpnrsgkalldvyvplgkisygqifgfygysg 420

```

```

OY 421 IDYNHEATSPGVGLMLNDPMGL 442
DB 421 idynheatsfgvglmndwmgl 442
RESULT 5
AAV75156
ID AAV75156 standard; Protein: 370 AA.
XX
XX AAV75156;
XX
XX 21-MAR-2000 (first entry)
XX
XX Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX antibacterial; gene therapy.
XX
XX Neisseria gonorrhoeae.
XX
XX W09957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
XX
XX 31-JUL-1998; 98US-0094869.
XX
XX 02-SEP-1998; 98US-0098994.
XX
XX 02-SEP-1998; 98US-0099062.
XX
XX 09-OCT-1998; 98US-0103749.
XX
XX 09-OCT-1998; 98US-0103794.
XX
XX 09-OCT-1998; 98US-0103796.
XX
XX 25-FEB-1999; 99US-0121528.
XX
XX (CHIRON ) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Petersen J, Piazza M, Rappuoli R, Ratcl G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX
XX WPI: 2000-062150/05.
XX N-PSDB: AA253918.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics
XX
XX Claim 2; Page 903; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941
XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of
XX the invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the
XX manufacture of medicaments for treating or preventing infection due to
XX Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
XX presence of Neisseria bacteria, or to raise antibodies. They may also
XX be used to screen for agonists or antagonists, which may themselves
XX have use as antibacterial agents. The polynucleotides of the invention
XX may also be used in gene therapy protocols.
XX
XX Sequence 370 AA:
SQ
Query Match 34.8%; Score 822; DB 21; Length 370;
Best Local Similarity 44.7%; Pred. No. 8.5e-71;
Matches 163; Conservative 68; Mismatches 110; Indels 24; Gaps 8;
OY 87 INCSALNDIMRLACVDTLVHGTPRAVIT-----TKKRSIRLDEETIMOTI-KGKPPVVQVE 140

```



```

PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0098994.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103794.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizze M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR N-PSDB; AA53920.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
XX Claim 2; Page 905; 1453pp; English.
XX
XX AA539015 to AA54536, AA54577 to AA54615, and AA74253 to AA75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA54537 to AA54576 and AA54616 to AA5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
XX Sequence 370 AA:
SQ

```

```

Query Match 34.5%; Score 815; DB 21; Length 370;
Best Local Similarity 44.7%; Pred. No. 4,1e-70;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

```

```

QY 87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKRSIRLDETIWQTI-KGKPQVYQOE 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 21 lqcaaltldvttlaacydrtifaaglpesagqegeskavlnlctvrsldkgeavtlvek 80

QY 141 TTDPITFLMGNENKGMILTKDAKOLEYAAKQFTPLSLSFDDLRNN-TPLWSSRPHNPYVLP 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 81 ggda1-----padsagetadlytprlsimydldkndlgllgyvrehnmpylnp 127

QY 200 IFMHGKPNRSPMTTPSHEAR-QFTPNFRAPELKFQVSVKVAEADMGTDSLMTFGYTQO 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 lwynspnyapapsprgtvtgvekfgqktraetklqvstfskslaedlfxtradiwfygtqr 187

QY 259 SHMOJFN-GKNSRPRFVNDYOPEIFLTQPVYSDLPMDGKVRMIGAVHNSGESAKLSR 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 188 sdwqlyngqirsafrntcykpeflftqpvkadlrfggrlmlnlgagfynqsnqgsrpsr 247

QY 318 SWNRAYLAGMEKMKLTVMPRIWGRIFKEGSGSQDDPDLDYGYGVDFLYOLENKS 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 248 swnrilyamagmewgklvtlprvwvrafdq-sgdk-ndnpdldadnyngydvklyrIndrq 305

QY 378 NISGTVRYNPNRSGKALQLDYVYPLGKGISGYFOIFGQYGSGLIDYNHBAISFGVGLMLN 437
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 nysvslitvmpkctgyaleaayclfpklklygvrtfhyggeslidygnhkqngi9lmln 365

QY 438 DMMGL 442
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 366 dldgl 370

```

```

RESULT 8
AA70629
ID AA70629 standard; Protein; 374 AA.
AC AA70629;
DE 18-JUL-2000 (first entry)
XX
XX Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
XX
XX BASB033: diagnosis, prophylaxis; treatment; antibacterial; vaccine;
XX Neisseria meningitidis infection.
XX
XX Neisseria meningitidis.
XX
XX WO200015801-A1.
XX
XX 23-MAR-2000.
XX
XX 09-SEP-1999; 99WO-EP06718.
XX
XX 14-SEP-1998; 98GB-0020003.
XX
XX (SMRK ) SMTTKLINE BECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI; 2000-271439/23.
DR N-PSDB; AA52134.
XX
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis infection -
PT meningitidis infection.
XX
XX Claim 4; Page 59; 93pp; English.
XX
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain H44/76. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
XX Sequence 374 AA:
SQ

```

```

Query Match 34.5%; Score 815; DB 21; Length 374;
Best Local Similarity 44.7%; Pred. No. 4,1e-70;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

```

```

QY 87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKRSIRLDETIWQTI-KGKPQVYQOE 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 25 lqcaaltldvttlaacydrtifaaglpesagqegeskavlnlctvrsldkgeavtlvek 84

QY 141 TTDPITFLMGNENKGMILTKDAKOLEYAAKQFTPLSLSFDDLRNN-TPLWSSRPHNPYVLP 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 ggda1-----padsagetadlytprlsimydldkndlgllgyvrehnmpylnp 131

QY 200 IFMHGKPNRSPMTTPSHEAR-QFTPNFRAPELKFQVSVKVAEADMGTDSLMTFGYTQO 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 132 lwynspnyapapsprgtvtgvekfgqktraetklqvstfskslaedlfxtradiwfygtqr 191

QY 259 SHMOJFN-GKNSRPRFVNDYOPEIFLTQPVYSDLPMDGKVRMIGAVHNSGESAKLSR 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 192 sdwqlyngqirsafrntcykpeflftqpvkadlrfggrlmlnlgagfynqsnqgsrpsr 251

QY 318 SWNRAYLAGMEKMKLTVMPRIWGRIFKEGSGSQDDPDLDYGYGVDFLYOLENKS 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 swnrilyamagmewgklvtlprvwvrafdq-sgdk-ndnpdldadnyngydvklyrIndrq 309

QY 378 NISGTVRYNPNRSGKALQLDYVYPLGKGISGYFOIFGQYGSGLIDYNHBAISFGVGLMLN 437

```

Db	310	nvysvrlrnpktyggaiaeaayflrplkgrlkgyvrrfngygeslidyhmkngigslmfn	369
Qy	438	DWGL 442	
Db	370	dldgl 374	
RESULT 9			
ID	AAV70628	standard; Protein; 375 AA.	
XX	AAV70628;		
XX	18-JUL-2000	(first entry)	
DE	Neisseria meningitidis serogroup B strain ATCC13090 BASB03 protein.		
XX	BASB03; diagnosis, prophylaxis; treatment; antibacterial; vaccine;		
KW	Neisseria meningitidis infection.		
XX	Neisseria meningitidis.		
XX	WO200015801-A1.		
XX	23-MAR-2000.		
XX	09-SEP-1999;	99WO-EP06718.	
PF	14-SEP-1998;	98GB-0020003.	
PR	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.		
XX	Ruelle J;		
XX	WPI: 2000-271439/23.		
DR	N-PSDB; AA252133.		
XX	Isolated BASB03 polypeptides and polynucleotides of Neisseria		
PT	meningitidis, useful for diagnosis, prophylaxis and treatment of N.		
PT	meningitidis infection -		
PS	Claim 4; Page 58; 93pp; English.		
XX			
CC	The present sequence is a BASB03 protein from		
CC	Neisseria meningitidis serogroup B strain ATCC13090. The protein		
CC	shows homology to the Klebsiella pneumoniae outer membrane		
CC	phospholipase A. The present sequence is useful for diagnosis,		
CC	prophylaxis and treatment of N. meningitidis infection. It may also be		
CC	used for the discovery and development of antibacterial compounds and		
CC	in vaccine compositions.		
XX			
SQ	Sequence 375 AA;		
Qy	Query Match	34.4%; Score 812; DB 21; Length 375;	
	Best Local Similarity	44.4%; Pred. No. 8.1e-70;	
	Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;		
Db	87	INCSLLNDIMRLACIDYLHGETPAVI-----KTKRSIRLDETWTOT-KGKPOVYTOE	140
Db	26	lqcaaltlvtrlycydrflfaaqlpsasagqegqeskavlnlletvrssldkgeavilvek	85
Qy	141	TDPFLFMNGENKGMJTKKDAQLEVAARQFPLSLSPFLDRNN-TPLMSSRPHNPMYILP	199
Db	86	ggdaa-----padasagetadilyrplslmyddkrdldglllgvhehnpmylmp	132
Qy	200	IFMHGKPNRSPTPSHEAR-OFTPNEFRAPBLKFOVSVKVAEDLWGSDSLWFGYTOO	258
Db	133	lwynsppryapsprctgtygekfgqgkrraetklyvstfkskaedlftktradrwfygtyqr	192
Qy	259	SHWQIFN-GKNSRPRVHDYQPEIFLTOPVYSDELPMWCKVRMIGAVHSHNGESAKLSR	317

```

Db      193 scdwqlynggrkbsapfrindcykpeifltqpvkcdlpefgslmlagfvthqnsngsrpsar 252
Qy      318 SNNRAYLWAGMEWKNLTVPRIWGRIFEGSGSQDPDNDIIDDYGYGQVRFYLQLENKS 377
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      253 swrilyamagmewgkltiviprvwvrafdq -sgdk -ndnpdiadymgysgdvklqyrlndrq 310
Qy      378 NISGVTRYPRBSGKALQLDYVYPLGKGISGFQIFQIGSGSLIDYVNHKATSFQVGLMIN 437
        | : : ||||| : : ||||| : : ||||| : : ||||| : : ||||| : : ||||| : :
Db      311 nuyssvlytrnpkctgygaleaayftplkqklkgyvrgfhygselidyhkhqnglqldmfn 370
Qy      438 DWMGL 442
        | :
Db      371 dldgi 375

RESULT 10
AAM98871
ID      AAM98871 standard; Protein; 355 AA.
XX
XX      AAM98871;
XX
XX      31-MAR-1999 (first entry)
XX
XX      H. pylori GHPO 1723 protein.
DE
XX      GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX      peptic ulcer disease.
XX
XX      Helicobacter pylori.
OS
XX      WO9843478-A1.
XX
XX      08-OCT-1998.
XX
XX      01-APR-1998; 98MO-US06371.
XX
XX      29-JUL-1997; 97US-0902615.
XX      01-APR-1997; 97US-0833457.
XX      24-JUN-1997; 97US-0881227.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
PA      (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PI
PI      Al-Garawi A, Kleantous H, Miller C, Oomen RP, Tomb J;
XX
XX      WPI; 1998-542293/46.
XX      DR      N-PSDB; AAX14590.
XX
XX      New isolated Helicobacter polynucleotides - used to develop products
XX      for the diagnosis, prevention and treatment of Helicobacter
XX      infections and gastrointestinal diseases
XX
XX      Claim 8; Page 1976-1977; 2054pp; English.
XX
XX      This sequence represents a Helicobacter pylori GHPO protein of the
XX      invention. The polypeptides can be used for preventing or treating
XX      Helicobacter infections, and gastroduodenal diseases associated with
XX      these infections, including acute, chronic, and atrophic gastritis, and
XX      peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
XX      used for the production of antibodies. The products can also be used for
XX      detection and diagnosis.
XX
XX      Sequence 355 AA;

Query Match 10.4%; Score 246.5; DB 19; Length 355;
Best Local Similarity 25.1%; Pred. No. 3.2e-15;
Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

Qy      157 KKDAGOLEAAQOFPLSLSPDLDDNNPTLWMSRRPNRPYVLPDIFMHGKPNRSPPTPSHE 216
        || : || : || : || : || : || : || : || : || : || : || : || : || :
Db      69 kkyllmmudylqetyflpbfynsf-----tpilqwyhpnlnp----- 102

```

```

QY 217 AROFTPNEFARPELKFQVSVKKAEDLMGTDSDLMFGYTQOSHQIFNGKNSRPFRRVD 276
DB 103 ---ygrnef-----kfqlsfvrvfthlwtckgtlylaytqtdwftqlyndpsapmrtn 154
QY 277 YQPEIFLTQPVYSDLPWDCKV---RWIGMAVHHSNG-ESAKLSRSWNAAYLMAGMEKN 332
DB 155 fmpelllyvypI-nfkpfqgkignfseilwqhlisngvgaqcyqpfnk-----egnpeng 209
QY 333 LVVMPRI-----WGRIFKSGSGSP-----DDNDP 357
DB 210 fpgqpylrvkdyngqkdvrvwgscsvsaqgrpvfrlvwekgjlklmwaywpyrvpydqsnp 269
QY 358 IIDVYGYGVDFVFLY-----OLENKSNIISGTVRYNPRSGKALQLDVYVPLGKIGSYF 410
DB 270 Ildymygnakldyrrgrhfeqlqlydftqwyrd--rwngafrlgyttrlnptv 327
QY 411 QIFQGYGOSLIDYNHEATSFVGGLMLN 437
DB 328 gwfngygdglveydvfsnrlygvgrln 354

RESULT 11
AAV10960
ID AAV10960 standard; Protein; 356 AA.
XX
AC AAV10960;
XX
DT 08-JUN-1999 (first entry)
XX
DE H. pylori ORF 07ap8060L_5083193_f3_8 cell envelope protein.
XX
KM Vaccine; probe: diagnostic; ORF; cell envelope protein;
XX secreted protein; cellular protein.
XX
OS Helicobacter pylori.
XX
PN WO9818323-A1.
PD 07-MAY-1998.
XX
PF 28-OCT-1997; 97WO-US19575.
XX
PR 14-JUL-1997; 97US-0891928.
PR 28-OCT-1996; 96US-0738150.
PR 06-DEC-1996; 96US-0759739.
XX
PA (ASTR ) ASTRA AB.
XX
PI Alm RA, Smith D;
XX
DR WPI; 1998-271811/24.
DR N-PSDB; AAX30427.
XX
PT Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
PT infections
XX
PS Claims 27, 31; Page 160-161; 279pp; English.
XX
CC Recombinant or substantially pure preparations of H. pylori polypeptides
XX are disclosed, together with the nucleic acids encoding them. In all,
XX 73 ORFs are shown. The proteins are variously cell envelope proteins,
XX secreted proteins or other cellular proteins. Vaccines containing the
XX nucleic acids or proteins are claimed, as are probes containing at least
XX 8 nucleotides from the nucleic acid sequences. The vaccines are useful
XX for treating or reducing the risk of H. pylori infections, and the
XX probes can be used diagnostically for detecting the presence of
XX Helicobacter in a sample. The products are also of use in screening
XX for compounds having the ability to interfere with the H. pylori life
XX cycle or to inhibit H. pylori infection.
XX
SQ Sequence 356 AA;

```

```

Query Match 10.1%; Score 239.5; DB 19; Length 356;
Best Local Similarity 24.8%; Pred. No. 1.5e-14;
Matches 82; Conservative 43; Mismatches 11; Indels 95; Gaps 12;

QY 157 KKDAAQOLEVAAQOFPLISFDLDNRNPTLMSSRRHNPVYVPIFMHGKPNRSPNTPSHE 216
DB 70 Kkylmmmdylgtyflrfynst-----tpilqwyhplnlp----- 103
QY 217 AROFTPNEFARPELKFQVSVKKAEDLMGTDSDLMFGYTQOSHQIFNGKNSRPFRRVD 276
DB 104 ---ygrnef-----kfqlsfvrvfthlwtckgtlylaytqtdwftqlyndpsapmrtn 155
QY 277 YQPEIFLTQPVYSDLPWDCKV---RWIGMAVHHSNG-ESAKLSRSWNA----- 322
DB 156 fmpelllyvypI-nfkpfqgkignfseilwqhlisngvgaqcyqpfnkengnpengfpg 214
QY 323 -----YLMAGMEKN--NLVMPRIWGRIFKSGSGSPDP- 354
DB 215 pvlvkkdyngqkdvrvwgscsvxxgnxlcfvlwkegjlklmwaywpyv-----pydq 266
QY 355 -NPDLIDVYGYGVDFVFLY-----OLENKSNIISGTVRYNPRSGKALQLDVYVPLGKGI 406
DB 267 snpqldymygnakldyrrgrhfeqlqlydftqwyrd--rwngafrlgyttrlnptv 324
QY 407 SGYFOIFQGYGOSLIDYNHEATSFVGGLMLN 437
DB 325 giyaqwfngygdglveydvfsnrlygvgrln 355

RESULT 12
AAW20760
ID AAW20760 standard; Protein; 253 AA.
XX
AC AAW20760;
XX
DT 15-JUL-1997 (first entry)
XX
DE H. pylori outer membrane protein, 07ap80601orf8.
XX
KM Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KM binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KM duodenal ulcer disease; chronic gastritis; diagnosis; envelope;
KM outer membrane.
XX
OS Helicobacter pylori.
XX
PN WO9640893-A1.
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US09122.
XX
PR 01-APR-1996; 96US-0630405.
PR 07-JUN-1995; 95US-0487032.
XX
PA (ASTR ) ASTRA AB.
XX
PI Berglindh OT, Smith D, Møllgaard B;
XX
DR WPI; 1997-052306/05.
DR N-PSDB; AAT68013.
XX
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
XX
PS Claim 56; Page 1172-1173; 1481pp; English.
XX
CC The present sequence is a Helicobacter pylori outer membrane protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds, useful
CC as potential H. pylori life cycle activators or inhibitors. The genomic

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:10:38 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 2360
Sequence: 1 MKVSLSLTLTSLISCFALIA.....YNHEATSFQVGLMINDMGL 442

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq_1101.*
1: /SIDSB/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
2: /SIDSB/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
3: /SIDSB/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
4: /SIDSB/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
5: /SIDSB/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
6: /SIDSB/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
7: /SIDSB/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
8: /SIDSB/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
9: /SIDSB/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
10: /SIDSB/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
11: /SIDSB/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
12: /SIDSB/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
13: /SIDSB/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
14: /SIDSB/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
15: /SIDSB/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
16: /SIDSB/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
17: /SIDSB/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
18: /SIDSB/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
19: /SIDSB/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
20: /SIDSB/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
21: /SIDSB/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
22: /SIDSB/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2360	100.0	442	21	BASB034 amino acid
2	2354	99.7	442	21	BASB034 amino acid
3	2352	99.7	442	21	BASB034 amino acid
4	2330	98.7	442	21	BASB034 amino acid
5	821	34.8	370	21	Neisseria meningit
6	814	34.5	370	21	Neisseria meningit
7	814	34.5	370	21	Neisseria meningit
8	814	34.5	374	21	Neisseria meningit
9	811	34.4	375	21	Neisseria meningit
10	246.5	10.4	355	19	H. pylori GHP0 172
11	239.5	10.1	356	19	H. pylori ORF 07ap

12	156.5	6.6	253	18	AAW20760
13	117	5.0	1686	19	AAW70991
14	110.5	4.7	800	8	AAW70420
15	110.5	4.7	822	13	AAW26021
16	109	4.6	824	21	AAW23180
17	107	4.5	157	18	AAW20538
18	103	4.4	798	21	AAW40925
19	103	4.4	798	21	AAW90225
20	103	4.4	798	22	AAW83944
21	102.5	4.3	537	22	AAW91000
22	100.5	4.3	1726	18	AAW38756
23	98.5	4.2	467	20	AAW13378
24	98.5	4.2	467	21	AAW94860
25	98.5	4.2	467	22	AAW86401
26	98.5	4.2	467	22	AAW80246
27	96	4.1	888	22	AAW67051
28	95.5	4.0	502	22	AAW90031
29	95.5	4.0	516	22	AAW78876
30	95.5	4.0	682	17	AAW04359
31	95	4.0	1227	22	AAW81501
32	94	4.0	372	16	AAW69607
33	94	4.0	522	22	AAW95616
34	93.5	4.0	467	19	AAW75057
35	93	3.9	761	20	AAW99084
36	93	3.9	1717	22	AAW20498
37	92	3.9	600	22	AAW75092
38	92	3.9	1627	16	AAW67538
39	91.5	3.9	857	13	AAW29614
40	91.5	3.9	1024	18	AAW19604
41	91	3.9	564	21	AAW10457
42	91	3.9	564	21	AAW81952
43	90	3.8	522	22	AAW95513
44	90	3.8	790	22	AAW93045
45	90	3.8	841	20	AAW24318

ALIGNMENTS

RESULT 1
ID AAY85268 standard; Protein; 442 AA.
XX
AC AAY85268;
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #1.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WC2000J5802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPL; 2000-271440/23.
XX
PT N-PSDB; AAA10700.
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Fig 2; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderly, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
XX Sequence 442 AA;
SQ

Query Match 100.0%; Score 2360; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 4.5e-219;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSILSCFALIAIOQAQAVNPVAFYDEVRESENDLCODNELPIDVOSATOSA 60
DB 1 mkvslstltlsilscfaliaiqgaqavnpvafdevreSENDlgdneIdpIdvgsatgsa 60
QY 61 STDTPANPLDEHEPELYTALENKMTLINCASALNODIMRLACYDTLVHETPAVITKTKSI 120
DB 61 stdtpanpldehePElytalenkmtlincasalnOdImrlacydTLvHETpavItkTksi 120
QY 121 RLDETITQTIKGRPOVYIQTETDPIFLMGNEKGMILTKDAKOLEYAAKQFPLSLSPDL 180
DB 121 rldetiTwitkgRPOViyqTETDPIflmgNEKgmILtkdAKOLEyAAKQfPLslSPdl 180
QY 121 rldetiWtIkgRPOViyqTETDPIfLmgNEKgmILtkdAKOLEyAAKQfPLslSPdl 180
DB 121 rldetiWtIkgRPOViyqTETDPIfLmgNEKgmILtkdAKOLEyAAKQfPLslSPdl 180
QY 181 RNNTPPLWSSRPHPNPYVLPPIFMHGKPNRSPMTPSHAKQFTPNFERABELKFOYSVKYA 240
DB 181 rnnTPPlwSSrPhpNPYvLPPIfMHgKPNrSPmTPshAKQfTPnFERABELKfOysvKya 240
QY 241 AEDLMGTSDLMFGYTOOSHQIFRNGKNSRPRVNDYOPETFLQPVYSDLPMDGKVM 300
DB 241 aedlmgTsdLMfgyTOOSHqIFrNgKNSrPRvNDyOPETfLQpVYsDLPmDgKvM 300
QY 241 aedlmgTsdLwfgYtOOSHqIfrNgKnsrPRfVndYOpETfLpVysdLpWdgkvM 300
DB 241 aedlmgTsdLwfgYtqgshwqIngrKnsrpfVndYOpETfLpVysdLpWdgkvM 300
QY 301 GMAVHSHNGSASAKRSRMRNAYLMAGMEKMLTFVMPRIKGRIFEGSGSQDDNDPDL 360
DB 301 gmaVhshngsAsakRSrMRnAYlMAGmEkMLTFvMPRIKGRIfEGsgsqDDndPdl 360
QY 361 YVYGVDFVFLYQLENKNSISGTVRYNPRSGALQLDVYVPLGKGISGYPOLFQYQGS 420
DB 361 yvYgvDfVflYqLenKNSISgTVryNpRsgALQldVYvPlGkGISgYpOLFqYqGS 420
QY 421 IDYHHEATSPFVGLMDNDMGL 442
DB 421 idYhHEATsPfVglMDndMgl 442

RESULT 2
AAV85271
ID AAV85271 standard; Protein; 442 AA.
XX
AC AAV85271;
XX

DT 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #4.
DE
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
OS
PN WC0200015802-A1.
XX
XX 23-MAR-2000.
PD
XX
XX 14-SEP-1999; 99WO-EP06781.
PF
XX
XX 14-SEP-1998; 98GB-0020002.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
PI Ruelle J;
XX
XX WPI: 2000-271440/23.
DR
XX N-PSDB: AAA10703.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
XX
PS Claim 3; Page 69; 106pp; English.
XX

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2969. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderly, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA;
SQ

Query Match 99.7%; Score 2354; DB 21; Length 442;
Best Local Similarity 99.5%; Pred. No. 1.7e-218;
Matches 440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSILSCFALIAIOQAQAVNPVAFYDEVRESENDLCODNELPIDVOSATOSA 60
DB 1 mkvslstltlsilscfaliaiqgaqavnpvafdevreSENDlgdneIdpIdvgsatgsa 60
QY 61 STDTPANPLDEHEPELYTALENKMTLINCASALNODIMRLACYDTLVHETPAVITKTKSI 120
DB 61 stdtpanpldehePElytalenkmtlincasalnOdImrlacydTLvHETpavItkTksi 120
QY 121 RLDETITQTIKGRPOVYIQTETDPIFLMGNEKGMILTKDAKOLEYAAKQFPLSLSPDL 180
DB 121 rldetiTwitkgRPOViyqTETDPIflmgNEKgmILtkdAKOLEyAAKQfPLslSPdl 180

PR 14-SEP-1998; 98GB-0020002.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Ruelle J;
 XX WPI: 2000-271440/23.
 DR N-PSDB: AAA10701.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 XX
 PS Claim 3; Page 67; 106pp: English.
 XX
 CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderly, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 CC
 XX Sequence 442 AA;
 XX
 SQ

Query Match 98.7%; Score 2330; DB 21; Length 442;
 Best Local Similarity 98.6%; Pred. No. 3.5e-216;
 Matches 436; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKVSI STLTLII SCFALIAIQAAVNPVAFVDEVRSENDLGDNLEPLDVOSATOSA 60
 DB 1 mkvslstltlilpcfallaigqagavnpvafvdevrskndlgqdnelllygvaqtgsa 60
 QY 61 STDTPANPLDEHEPELYTTALENKTMLINCASALNODIMRLACYDTLVHGEPYAVIKTKRSI 120
 DB 61 stdtanpldehepeelyttalenktmlincasalnodimrlacydtlvhgepaviktkrsi 120
 QY 121 RLDETIWITIGKPOVYIQTETDPFLFMNGNEKMTTKDAQOLEVAAQFTPLSIFPLD 180
 DB 121 rldeitiwiti gkpo vy iqtetdp flfmngnekmttkdaql e vaaqftplslstifld 180
 QY 181 RNNPFLMGSSRPNNPVYPIFMHGKPNRSPNTPSHEAKQFTPNERRAELKFOVSVYTKA 240
 DB 181 rnnpflmgssrpnnpvy pifmhgkpnrs pntpsheakqftpnerraelkfovsvytk a 240
 QY 241 AEDLMGTDSDLMFGYTOQSHWOIFNGKNSRPRVHYDYOPELFLQPVYSDLPWVGKXVMI 300
 DB 241 aedlmgt d s d l m f g y t o q s h w o i f n g k n s r p r v h y d y o p e l f l q p v y s d l p w v g k x v m i 300
 QY 301 GCGAVHNSNGSAKISRSNRAVYLMAGEMKNLTVMPRIWGRIRKESGSGDDNDPDLTD 360
 DB 301 gmgavhnsngsakisrsnraylmagewknltvmpriwgrirke sgsqddndpdltd 360
 QY 361 YYGVDVAFELVQLENKNSIGTVRNPRSGALQDLYVYLGGISYFQIPFGYGGSL 420
 DB 361 yygvdvafel vq lenkns ig tv r n p r s g a l q d l y v y l g g i s y f q i p f g y g g s l 420

QY 421 IDYNHEATFSGVGLMDNDMGL 442
 DB 421 idynheatsfgvglmndmgl 442

RESULT 5
 AAY75156
 ID AAY75156 standard; Protein; 370 AA.
 XX
 AC AAY75156;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 OS Neisseria gonorrhoeae.
 XX
 PN W0957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 DR WPI: 2000-062150/05.
 DR N-PSDB: AA253918.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 PS
 PS Claim 2; Page 903; 1453pp: English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 370 AA;
 XX

Query Match 34.8%; Score 821; DB 21; Length 370;
 Best Local Similarity 44.4%; Pred. No. 1.2e-70;
 Matches 162; Conservative 69; Mismatches 110; Indels 24; Gaps 8;

QY 87 INCASALNODIMRLACYDTLVHGEPYAVI----KTKRSIRLDETIWQTI-KGKPOVYIQE 140

[illegible]

XX AA253015 to AA254536, AA254537 to AA254615, and AA274253 to AA275941 represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides and polypeptides. AA254537 to AA254576 and AA254616 to AA254743 represent PCR primers used in the exemplification of the present invention. The polypeptides, the polynucleotides, antibodies and compositions of the invention can be used as vaccines, as diagnostic reagents, and as immunogenic compositions. The polypeptides can be used in the manufacture of medicaments for treating or preventing infection due to *Neisseria* bacteria (e.g. meningitis and septicaemia), to detect the presence of *Neisseria* bacteria, or to raise antibodies. They may also be used to screen for agonists or antagonists, which may themselves have use as antibacterial agents. The polynucleotides of the invention may also be used in gene therapy protocols.

sequence of *H. pylori* (ATCC 55679) was determined from overlapping cDNAs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely *H. pylori* antigens for vaccine development, the amino acid homologies predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from *H. pylori* by PCR amplification for recombinant polypeptide production, e.g. in *E. coli* hosts.

SQ Sequence 253 AA;

Query Match	6.6%	Score 156.5	DB 18	Length 253
Best Local Similarity	25.6%	Pred. No. 9.6e-07		
Matches 56; Conservative	26	Mismatches 72	Indels 65	Gaps 8

Oy	157	KKAKQLEELAAKKOFPRSLSDLDNRNTPRLASSRPHNPVLYLPEFMGKRPRRSNPESHE	216
Dd	72	kylmmmdylylclrylfrrhsfr-----crlfywghnllnp-----	105
Oy	217	AkofPnPFRARPELFQVSvYVKAAEDLMGTSDLMEGYTQOSHQJFENGKNSRPFEVD	276
Dd	106	---ygrnef-----kfgdsfrvrfrnhlkcyklslayqctnwfdqlyndorqsapmtmin	157
Oy	277	yOPeIFLTOPYVSDLPMWDGY--RMIGMGAVHNHSG--ESAKLSSRNRAAYLWAGMEMKN	332
Dd	158	fmpellyuyprl-nfkprfgsklgntselawlgshlnsgvgagdcyqrftn-----	204
Oy	333	LTMVRIRNGRIFFKESGSGSQ--PDNDPILDIYYGGDNRAF	369
Dd	205	-----kegnpenqdfpgprpvlykdnyngkdvwtm	231

RESULT	13
AAW70991	
ID	AAW70991 standard; Protein; 1686 AA

AC AAW70991;

DT 19-OCT-1998 (first entry)

DE Human class II p13 kinase-C2alpha.

KW Human; class II phosphoinositide lipid kinase; PI3 kinase

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX
XX

holio sapientis.

PN MO9032604-AZ
XX

30-JUL-1998.
PD
XX

PE 27-JAN-1998; 98WO-G800244.
XY

PR 28-JAN-1997; 97GB-0001652.
XX
XXPA (LUDW-) LUDWIG INST CANCER RES.
XX

PI Domain J, Waterfield MD;

DR WPI; 1998-427960/36.

XX

PT fragments - useful for, e.g. treatment of tumour cells where

XX

XX
XX

kinase due to the presence of a conserved C2 domain found in murine and Drosophila class II p13 kinases, its apparent lack of a p85 binding site and a substrate affinity to inositol lipids Ptdins and Ptdins(4)P. The protein has resistance to p13 kinase inhibitors Wortmannin and LY294002. Antibodies against the protein (optionally humanised), are used to identify class II p13 kinases. Antisense sequences, antibodies or dominant negative mutants of the p13-Ca1pala protein, are useful in human or veterinary medicine to block class II kinases. They can be used to treat tumour cells where the phenotype is associated with expression of p13-Ca1pala protein.

Sequence 1686 AA;

Query Match	5.08;	Score 117;	DB 19;	Length 1686;
Best Local Similarity	19.28;	Pred. NO. 0.13;		
Matches 86;	Conservative 67;	Mismatches 154;	Indels 140;	Gaps 21;

```
OY 45 ODNELPIIV-----OSAQOSASTDPANLDE-----HEPELYTTALNKNKTNLIN-----88
Db 522 eddehpvdlnkhlyajekpeckeamrphveelldsyhnqvaelqlenqhvavdqvikav 581
OY 89 ---CSALMODIMRLACYPDTLVHGETPAVIKTKRSLRDETWTQIKCKPQVYOETDPI 145
Db 582 rkicsald-gyettai-----teskklkravnlpis-----ktadvt 618
OY 146 FLMGNEKMGLTKKADOLEVAAKOTPLISFDDLRNTPLWSSRPHNMUV-----LP 199
Db 619 slfge-----dsrst-rgslnpenpnvyslntlaa 651
OY 200 IF---HGCKPRSPNPTSHKAQOTTPEHFAPBLKFOVSXKYVAADL---WCTSDDLW 255
Db 652 lydlrlrhaansgrtpdcagssksvkawetleqqifilf----aainglsnwswsnky 707
```

QY 253 FGYQOSHNIENKKN-SRPF---RVHDYQPEIFLTLQ-----PV-YSDLPMDGKYRMI 300

BU 108 YLCSISIL---H9KALIKPIQSKVGLYNILYLILKWDIELIPRISQIPRESVINIL108

```

QY 301 GNGVHSHNGESAKLSRWN-----RAYLMAGMEWKNLIVMPRIWGRIEK 343

```

Db 764 LtglInqssgsspsdnkqrkqpealgvksLf

QY 346 EGGSGQPPDDNPDI LDYGYGVDVRELYQLENN

QY 406 ISGYFQIFQGYGSLIDYNHEATSFQV 432
| | : | : : | :
871 kca-----b11a11baacclal 007

RESULT 14

ID	AA	PF0420	standard; protein; 800 AA.
1

AC AAP70420;

DT 20-JAN-1991 (first en

Sequence encoded by ce

Enzyme; cellotriose; c

OS Bacillus sp. No. 1139.

aa	FH	Key	Locatt
aa	FH	Key	Locatt

FT	Peptide	FT	Protein
11.130	11.130	11.130	11.130
31.180	31.180	31.180	31.180

XX
PN
TP62232386-A

XX 12-OCT-1987
PD

PF 02-APR-1987; 87JP-0076285.
 XX 02-APR-1986; 86JP-0076285.
 XX (RIKA) RIKAGAKU KENKYUSHO.
 PA WPI: 1987-325245/46.
 XX N-PSDB: AAN70651.
 DR
 XX DNA sequence coding cellulase gene - is derived from *Bacillus* sp. no.
 PT 1139z and can hydrolyse cello-triose and cello-tetraose.
 XX
 XX Disclosure: Fig 3-1 - 3-3; 11pp; Japanese.
 PS
 CC This cellulase hydrolyses cellobiose or cellobiose and does not
 CC hydrolyse cellobiose. Molecular weight is approx. 92k dalton and
 CC optimum pH is 9.
 XX
 SQ Sequence 800 AA:

Query Match 4.7%; Score 110.5; DB 8; Length 800;
 Best Local Similarity 20.0%; Pred. No. 0.16;
 Matches 96; Conservative 53; Mismatches 163; Indels 169; Gaps 25;

QY 9 TLTSLSCFAILLAIQQAQAVPNVAFVDEVRSEND---LGQDNELPIDVQASTQASTPTA 65
 DB 7 tkqlissalllv1-lls1ftalaegnrtrednfkhllgndvkrpseagallqgevqg 65
 QY 66 NPL-DEHEPELY-----TTALENKTMLINCAL-----NODIMRLACY-DTLVHGERT 111
 DB 66 mtlvqgheklqlrgmstgltqwfpeellndnaykalandwesnmrlamygengyasn 125
 QY 112 AVIKTKRSIRLDETI-----WQT-----IKGRPOVY 138
 DB 126 eliksrvtlkgldalendmyivdwhvnapgqprpyagaedffrdlaalypphilly 185
 QY 139 QETTPP-----IFLMGNEKMLTKKDAKOLEYAAKQPTPLSFDLDRN---NTPLM 187
 DB 186 elanepssnnnggaglpnneegmaavk-----eyadplvemlrdsgnaddnlllvgsppw 240
 QY 188 SSRP-----HNPMVLPPIF--MHGKPNRS--PMTPSHEAKQPTNPEFRAPELKFO 233
 DB 241 sgrpldaadnpiddhntmtythvfytgshaastesyppeetpnsnergymntlya----le 296
 QY 234 VSVKKAEDLMGT-----DSDLMPGYTQO-----SHWOIFNGKNS-----RP 271
 DB 297 ngavafate--wgtsgangdgapyfdeadvwleflnemlswanwsltn-knevsqaftr 353
 QY 272 FRVH-----DYOPEIFLTPQPVYSDLPW---D 294
 DB 354 felgksnatsldpdpdqvwppeelslsgyvrarlkynvye--ldtklykvlwdfnd 410
 QY 295 GKVRMIGMGAHVHNGES-----AKLSRSMNRAYLMAEMKNTLVPRING 341
 DB 411 gtkgfgfv-----ngdspvedvvleneagalklsgldasndvsegnynanarlsadgwg 464
 QY 342 R 342
 DB 465 k 465

RESULT 15
 ID AAR26021 standard; Protein; 822 AA.
 XX AAR26021;
 XX
 XX 02-FEB-1993 (first entry)
 XX
 DE Alkaline cellulase K-64.
 XX
 KW Mass production; recombinant.

XX OS *Bacillus* sp. KSM-64 (FERM P-10482).
 XX PN JP04190793-A.
 XX PD 09-JUL-1992.
 XX PF 26-NOV-1990; 90JP-0324401.
 XX PR 26-NOV-1990; 90JP-0324401.
 XX PA (KAOS) KAO CORP.
 XX DR WPI: 1992-280112/34.
 XX N-PSDB: AAO27180.
 XX
 PT Alkaline cellulose mass prodn. - using recombinant plasmid and
 PT microorganism contg. cellulase gene
 PS
 PS Disclosure: Fig 4; 14pp; Japanese.

Query Match 4.7%; Score 110.5; DB 13; Length 822;
 Best Local Similarity 20.0%; Pred. No. 0.17;
 Matches 96; Conservative 53; Mismatches 163; Indels 169; Gaps 25;

QY 9 TLTSLSCFAILLAIQQAQAVPNVAFVDEVRSEND---LGQDNELPIDVQASTQASTPTA 65
 DB 7 tkqlissalllv1-lls1ftalaegnrtrednfkhllgndvkrpseagallqgevqg 65
 QY 66 NPL-DEHEPELY-----TTALENKTMLINCAL-----NODIMRLACY-DTLVHGERT 111
 DB 66 mtlvqgheklqlrgmstgltqwfpeellndnaykalandwesnmrlamygengyasn 125
 QY 112 AVIKTKRSIRLDETI-----WQT-----IKGRPOVY 138
 DB 126 eliksrvtlkgldalendmyivdwhvnapgqprpyagaedffrdlaalypphilly 185
 QY 139 QETTPP-----IFLMGNEKMLTKKDAKOLEYAAKQPTPLSFDLDRN---NTPLM 187
 DB 186 elanepssnnnggaglpnneegmaavk-----eyadplvemlrdsgnaddnlllvgsppw 240
 QY 188 SSRP-----HNPMVLPPIF--MHGKPNRS--PMTPSHEAKQPTNPEFRAPELKFO 233
 DB 241 sgrpldaadnpiddhntmtythvfytgshaastesyppeetpnsnergymntlya----le 296
 QY 234 VSVKKAEDLMGT-----DSDLMPGYTQO-----SHWOIFNGKNS-----RP 271
 DB 297 ngavafate--wgtsgangdggyfdeadvwleflnemlswanwsltn-knevsqaftr 353
 QY 272 FRVH-----DYOPEIFLTPQPVYSDLPW---D 294
 DB 354 felgksnatsldpdpdqvwppeelslsgyvrarlkynvye--ldtklykvlwdfnd 410
 QY 295 GKVRMIGMGAHVHNGES-----AKLSRSMNRAYLMAEMKNTLVPRING 341
 DB 411 gtkgfgfv-----ngdspvedvvleneagalklsgldasndvsegnynanarlsadgwg 464
 QY 342 R 342
 DB 465 k 465

Search completed: November 30, 2001, 14:16:58
 Job time: 380 sec

1

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/961,083
3 FILING DATE:
4 CLASSIFICATION: 435
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER:
7 FILING DATE:
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Brookes, A. Anders
10 REGISTRATION NUMBER: 36,373
11 REFERENCE/DOCKET NUMBER: PB340P2
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (301) 309-8504
14 TELEFAX: (301) 309-8512
15 INFORMATION FOR SEQ ID NO: 2:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 666 amino acids
18 TYPE: amino acid
19 STRANDEDNESS: single
20 TOPOLOGY: linear
21 MOLECULE TYPE: protein
22
23 US-08-961-083-2

```

```

: ATTORNEY/AGENT INFORMATION:
:
: NAME: White, John P.
:
: REGISTRATION NUMBER: 28,678
:
: REFERENCE/DOCKET NUMBER: 39116-A
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (212) 278-0400
:
: TELEFAX: (212) 391-0525
:
: INFORMATION FOR SEQ ID NO: 6:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 985 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: - MOLECULE TYPE: protein
:
PCT-US96-03916-6

```

Query Match 3.7%; Score 87; DB 5; Length 985;
Best Local Similarity 19.4%; Pred. No. 6.1;
Matches 73; Conservative 47; Mismatches 130; Gaps 20

```

QY 48 ELPIIDVOASOTASOTANPIIDHEHEPELYT-TALENTTMIINCALNODIMRACVOTLV 106
Db 563 QTPSTVPRKTIILSSST-----EPALFTTQAGAEALFQTSAAEPDMRKQSTTHF 6133
QY 107 HGEPPAVIKTRKRSIRLDETTWOTIKGRPOVIYOB-TTDPIT--FLAMENKGMILTKDAQ 162
Db 614 FTQAPSTVPKAT-----QTSTEBEVLYTQSPREPEVFTKLAAPBEIQTQPSA-- 6622
QY 163 LEYAAKQTPILSLSFDDDRNTPILMSSRP-----HN-----PMYLP---IFMHK 2051
Db 663 --APEVYTRSSSTMPETAOSTPLASONPTSSGCTHNTERTPTVQTPHTOKLYENK 719
QY 206 PNRSPNPS--HE--AKQFPR-NEFAPLEKPOVSUKVA-----AEDLIGTD 248
Db 720 TLFSEPTVYSEHEMSTAESOTPLLDVAKIVEKKSNDGEVATATCVSYKSPRYVETNKKVD 779
QY 249 -----SDLMFGYQOOSHWOJFNKNS-----RPF----- 2722
Db 780 LVDMVDEISGNSPAGVFNSENKMK--QKOLYRTDERTSYQVLMCLSCJSHSPREPLYFD 8366
QY 273 -----RVHNDQPELFLT---QPVYSDLPMDGKY-----RMIGMAVHNS 3091
Db 837 TSLIAREKDIPBELYFTSDPOTACTITLPSGVPRFEMSNNVSLDEYLTATTVVSHTA 8966
QY 310 GE-----SAKLSRSW 319
Db 897 GOSTVWKSSSARAAGEAW 912

```

RESULT 13
PCT-US96-03916-66
Sequence 66: APPLICATION PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND US95 THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 371-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO.: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03916-66

Query Match	3.7%;	Score 87;	DB 5;	Length 985;
Best Local Similarity	19.4%;	Pred. No. 6.1;		
Matches	73;	Conservative	47;	Mismatches 126;
			Indels	130;
			Gaps	20;

QY	48	ELPFDVOSATOSATDPANLDEHEPELYT-TALENTKTMILNCSALMODIMLACDYTLV	106
Db	563	QTBSTVEKTLTISST-----EPALFTRTOSAGTEATQTSNABPDTRMQSTETHF	613
QY	107	HGETPAVAKTKRSIRLDEITWOTIKGAPVUYOE-TTPDI--FLMGNEKMLTKDKAKO	162
Db	614	FTQAPSTVPKAT-----QFTSTDEPVLTGOSPREPEVPTRTIGABEITQTPSA--	662
QY	163	LEYAKOFTPLSLSFDDRRNTPMLSSRP-----HN-----PMVLP----IFMHCK	205
Db	663	---APEVYTRSSIMTPEASTPLASONPTSSGCTHTEPRTPYVQTPHTOKLYTENK	719
QY	206	PNRSPNPS--HE--AKOFP--NEPARALEKPVYSKYVA-----AEDLMGTD	248
Db	720	TLSFPTVYSEHEHMSHTESOTPLLDVAKIVEYKTSNDGEVTAATGVSVYKSPYRVEITMKWD	779
QY	249	-----SDLMWGYQOOSHWOJFNCKNS-----RPF-----	272
Db	780	LVDVMDSELGNSPAGVFNSENKMK--QKOLYRYRTDORTSVQMLCLSCSTSHSPEPCLFD	836
QY	273	-----RVHNDQPELFTL---QPVYSDLPMDGKY-----RMIGGAVHHNS	309
Db	837	TSLIAREKDIAPLELYFTSDQATCTTLPSGVVPREEMSLNNVSLPEVLYLTATTVVSHTA	896
QY	310	GE-----SAKLSRSM	319
Db	897	GQSTVWKSSARAGEAW	912

RESULT 14
US-08-961-083-2
Sequence 2, Application US/08961083
Patent No. 6158469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
City: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

Best Local Similarity 21.1%; Pred. No. 2.7; Matches 75; Conservative 51; Mismatches 110; Indels 119; Gaps 23;

```

QY 105 LVHGETPAVITKRSIRLEETIMQITKGPQVYQETTPIFLMG-----NEK---GMLTK 157
DB 24 LHPALSTIYNT-----LSTTESLTSSNKTLY---SPGSIFEVGFRRNSRMYLDGMWK 75

QY 158 K-DAKOLEYAAKQFPLSLSFDF-----LDRNNTPLMSSRRPHNMYVLPFIMHCK 205
DB 76 KYSDRIYVAVARNRDNPLNSAIGTLAKISGNVLDDHSNKPVMWTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAKQFT-----PNEFRAPELKQVSVKAKAEDIM 245
DB 126 -ERSPVAVELLANGNEVMDSSNINDASEYLMQSFDPYDTLLPEMKIGYLNKLT----- 177
QY 246 GUDSDLMFCYITQOSHWOITNGKNSRFRVHDQ-----PELFLNQ---PYVSDLPMDCK 296
DB 178 GANRFL-----TSWRSSDDPSGNE---SYKLETOSLPEFYLSPENPMHRSFGWNG- 226
QY 297 VNMIGMGAVH-----HSGNESAKLSRSWNRAY-----LMAGMEKMLTVP--RI 339
DB 227 IFPSGIPEDQKLSYMYNFIENNEEVAYFRMTNNSFYSLRLTILSGYFORLTYPSIRI 286
QY 340 WGRIFKESGSGOPD---NPDILDYGYGVRELYOLENKSNISETVR--YNPRS 389
DB 287 WNRFWSSPYDRCDYFIMCGP-----YAYCDV-----NTSPVCNCIOGFNPRN 329

```

RESULT 11

US-08-633-879C-2
Sequence 2, Application US/08633879C
Patent No. 5928922

```

GENERAL INFORMATION:
APPLICANT: Kivirikko, Karl I.
APPLICANT: Pihlajaniemi, Taina
APPLICANT: Heilaakoski, Tarja I.
APPLICANT: Annunen, Pia P.
APPLICANT: Nissi, Riitta K.
APPLICANT: No. 5928922e1ainen, Minna K.
TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
TITLE OF INVENTION: METHODS FOR PRODUCING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633.879C
FILING DATE: 10-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0041-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNTE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids

```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-633-879C-2

Query Match
Best Local Similarity 3.7%; Score 87; DB 2; Length 537;
Matches 98; Conservative 53; Mismatches 146; Indels 170; Gaps 25;

```

QY 1 MKVSLTLTLSTLISCPAILAIOQAAPVPVAVDEVRSENDLGQDNEPVIDQSA---- 56
DB 1 MKQVYLVLV-LMSWFGVLSWQAEEFTSIGHTDILYAEKDLVQSLKEVILVEBAKIAK 59
QY 57 -----TOSTASTD-----TANPLDEHE-----PELYTALENKT--MLINCS 90
DB 60 IKSMASKMEALTSRSADEGYLAHFNVAKYKRLNTDMPALGDVLVDASAGFVNLIS 119
QY 91 ALNQ-----DIMRLA-CY---DPLVHGETPAVITKRSIRLEDTI----- 126
DB 120 VQRQFPTEDESGAARALKRLDQTYKLPDITSRGELPET-KYQMLSVDDCFGIGRSA 178
QY 127 -----WQTIKGPQVYQETTPIFLMGNEKGMLTRK-DAKOLEYAAKQFPLSLSFDL 179
DB 179 YNEGDIYHTVLMMEQVLKQ-----LDAGEEATVTKSLVDLYSAVFQGLDHLRAVEL 231
QY 180 DRNNTPLMSSRRPHNMYVLPFIMHGRPNRSPNPSHEAK-----QTPNEFRAPELK 232
DB 232 TRRLSL-----DPSHERAGNLRYFERLLLEERGRSLSN 266
QY 233 QVSVKVKAAEDLWGTDSDLMFGYTOOSHWOIFNGKNSRPRVHDYQPEIFLQPYVSDLP 292
DB 267 QIDAGLATQENLY-----ERP---TDIPE-----RDVYESL- 295
QY 293 WDGKVMIGMGAVHNSGESAKLS-RSWNRAYLM--AGMEKMLTVPRIWGRIFREGSG 349
DB 296 -----ORGEVKVLTIPRQKLCFRYHNGNVPOLLALP-----FRE--- 331
QY 350 SQPDNDPDLIDYGYGVRELYOLENKSNI-----GYRYVPRBG 390
DB 332 EDEMDSPIHIVRY---DVMSDEIERIKETAKPKLARATVR-DPKTG 374

```

RESULT 12

PCT-US96-03916-6
Sequence 6, Application PC/TUS9603916
General Information:

```

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993

```

RESULT 9
US-08-478-882-2
; Sequence 2, Application US/08478882
; Patent NO. 5885575

```

Query Match 3.8%; Score 89.5; DB 2; Length 503;
Best Local Similarity 18.4%; Pred No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

OY 40 ENLDGQDELPIDVOSATQASASTDPANPLDEHEPELYTALLENKTMILNCALNODIMRL 99
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 169 QGDIDASNITIDEDVSS-----GSTIEKSTPEGLIHTDLPT--SQPGDRD---- 212
OY 100 ACYDPLVHGTEPAVYIKTKRSI-----RLDEITNQTICKPQVIYQETDPLFLMGNENG 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 213 --DAFFIGSTLATIATTPWVSANHTKQNERQWMPHISNPVLLQTYT----- 258
OY 154 MLTKKDAKOLEFAAQKFTPLSLSPFDLRNNTPL---WSSRPHNPMYVLPIMHGKPNRS 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 259 -----RMT-----DIDRSTSAHGEMTQEQPPF----- 283
OY 210 PNTPSHEAKQFTPNNEPRABELKFQVSVKKAEDLW-----GTDSDLMFGTYQOS 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 284 -NNHEQOEETPRH-----ATSTWADPNSTTEEAATQKEKWF-----EN 322
OY 260 HMOINGKN-----SRPRVHDYOPEIFLTQPVYSDDLW---DGKVRMIGMG 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 323 EMO---GKPPPPSEDSHTTEGTTASAHNNHPSQRRWTTQSOEDVSWTDFDPLSHPMGOG 379
OY 304 AVHHSNGESA 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 380 HOTESKGHSS 389

RESULT 10
US-07-717-331F-2
; Sequence 2, Application US/07717331F
; Patent No. 5484305
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07717, 331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-717-331F-2

```

Db 136 DLSAKNLNLTLSQSEVSLDLEFIETNFSDKNOLKLPDGSVVVANE---SYTVSVK 191
QY 236 VKYKAEDLWGTSDLMFGYTOOSHQOIFNGKNSRPFVHDYOPEI-FLTQPYSDL-PW 293
Db 192 ATLQKLKVLVYSNADSVGSAIPVYSLNGKTONDESEFPKSNINFAKNYNNALNPF 251
QY 294 DGVNRIGMGAVIHS--NGESAK-----LSRSMNRAYLAGMEKMLTVMPRIWGRIEK 345
Db 252 EAQOYLVGQGFNOKYVNMADVDKNDINNHIEQTFNVAKITA-----TLGKAFK 300
QY 346 ---EGSGSOP-----DDNPOLDY--GYGDV-----RFLYQL-- 373
Db 301 QGFCHKNGOFLSLKVLKSLGSLNNEFKOLFVYVRLGDPVSDLIQSSQSSNKKTYQLLF 360
QY 374 ENKSNI 379
Db 361 ENKTTI 366

RESULT 7

US-07-946-497-2
; Sequence 2, Application US/07946497
; Patent No. 5506119

; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT C044 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-946-497-2

Query Match 3.8%; Score 89.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

QY 40 ENDLGDNELPIDVOSATQSASTDTANPLDEHPELYTTALEKTKMLINSALNODIMRL 99
Db 169 QEDIDASNITIDEVSS-----GSTIEKSPREGIILHTDLP---SQPTGDRD---- 212

QY 100 ACYDTLVHGETPAVVIKTKRSI-----RIDETIMQTIKGRPOVIYQETTDPIFLMGNEKG 153
Db 213 ---DAFFIGSTLAIATTPVWASHTKONOROTQWNPISHNPEVLQTTT----- 258
QY 154 MLTKKAKOLEYAKQOTPLSLSFDDRRNTPL---WSSRPHNPVVLITFMHGRNRS 209
Db 259 -----RMT-----DIDRNSTAHGEMWTQEPQPPF----- 283
QY 210 PNTPSHAKQFTNPFAPRLKFOVSVKRAEDLW-----GTDSDLMEGYTOQS 259
Db 284 -NNHEYODEEETH-----ATSTWADPNSTTEAATQAEKWF-----EN 322
QY 260 HMOIFNGKN-----SRPFVHDYQPEIFLTQPYVSDLPW---DGKVRMIGMG 303
Db 323 EMQ---GKNPPTSEDSHVTGTTASAHNNHPSQRMRTTQSQEDVSWTDFDPLISHPMGQG 379
QY 304 AVHNSGESA 313
Db 380 HOTESKGHSS 389

RESULT 8

US-08-483-322-2
; Sequence 2, Application US/08483322
; Patent No. 5760178

; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT C044 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,322
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-483-322-2

Query Match 3.8%; Score 89.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:14:03 ; Search time 36.79 Seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHEATSGVGLMLNDWGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	4.3	1726	US-08-609-049A-30	Sequence 30, Appl
2	101.5	4.3	1726	US-09-170-996-30	Sequence 30, Appl
3	100.5	4.3	1658	US-08-609-049A-13	Sequence 13, Appl
4	100.5	4.3	1658	US-09-170-996-13	Sequence 13, Appl
5	95.5	4.0	682	US-08-481-435-6	Sequence 6, Appl
6	91.5	3.9	1024	US-09-091-117-5	Sequence 5, Appl
7	89.5	3.8	503	US-07-946-497-2	Sequence 2, Appl
8	89.5	3.8	503	US-08-483-322-2	Sequence 2, Appl
9	89.5	3.8	503	US-08-478-882-2	Sequence 2, Appl
10	89.5	3.8	857	US-07-717-331F-2	Sequence 2, Appl
11	87	3.7	537	US-08-633-879C-2	Sequence 2, Appl
12	87	3.7	985	PCT-US96-03916-6	Sequence 6, Appl
13	87	3.7	985	PCT-US96-03916-66	Sequence 66, Appl
14	86.5	3.7	666	US-08-961-083-2	Sequence 2, Appl
15	86	3.6	781	US-08-373-134D-2	Sequence 2, Appl
16	86	3.6	781	US-09-114-637-2	Sequence 2, Appl
17	86	3.6	1627	US-07-665-792E-9	Sequence 9, Appl
18	84	3.6	522	RE34606-6	Patent No. RE34, 60
19	84	3.6	816	US-07-731-157A-4	Sequence 4, Appl
20	84	3.6	816	US-08-229-444B-2	Sequence 2, Appl
21	84	3.6	816	US-08-541-780-4	Sequence 4, Appl
22	83.5	3.5	774	US-08-902-632-2	Sequence 2, Appl
23	83.5	3.5	774	US-09-073-354-1	Sequence 1, Appl
24	83.5	3.5	774	US-08-656-005A-1	Sequence 1, Appl
25	83.5	3.5	774	US-09-073-259-1	Sequence 1, Appl
26	83.5	3.5	774	US-09-363-095-1	Sequence 1, Appl
27	83.5	3.5	774	US-09-418-027-1	Sequence 1, Appl

28	83	3.5	355	1	US-07-946-497-5	Sequence 5, Appl
29	83	3.5	355	1	US-08-483-322-5	Sequence 5, Appl
30	83	3.5	355	1	US-08-478-882-5	Sequence 5, Appl
31	82.5	3.5	535	2	US-08-564-972-1	Sequence 1, Appl
32	82.5	3.5	535	4	US-09-171-969-9	Sequence 9, Appl
33	82.5	3.5	657	4	US-09-306-593-2	Sequence 2, Appl
34	82.5	3.5	663	1	US-08-441-139-7	Sequence 7, Appl
35	82	3.5	334	2	US-08-359-850-4	Sequence 4, Appl
36	82	3.5	527	2	US-08-592-126-145	Sequence 145, App
37	82	3.5	527	2	US-08-687-080-48	Sequence 48, Appl
38	82	3.5	535	2	US-08-633-879C-4	Sequence 4, Appl
39	82	3.5	797	4	US-09-086-912-2	Sequence 2, Appl
40	82	3.5	1022	3	US-08-772-270A-2	Sequence 2, Appl
41	82	3.5	1088	3	US-08-633-768A-1	Sequence 1, Appl
42	81	3.4	320	2	US-08-245-511-4	Sequence 4, Appl
43	81	3.4	320	2	US-08-600-993A-4	Sequence 4, Appl
44	80.5	3.4	354	1	US-07-946-497-4	Sequence 4, Appl
45	80.5	3.4	354	1	US-08-483-322-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-609-049A-30
; Sequence 30, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-609-049A-30

Query Match 4.3%; Score 101.5; DB 2; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.47;
Matches 98; Conservative 80; Mismatches 189; Indels 161; Gaps 21;
QY 1 MKVSLSTLTSLSCFAILAIAQAAVNPVAFVD-----EVRSENDLQGD 46
DB 465 VKVSIIEGLQLPVTTCDSVSTVEIIIMQALSWVHDDLNQVDVGSVILKVCQEEVLQN 524

QY 47 NELPIDVQSATQSSTDPANPLDEHEPELYTTALENKTMLINCNSALNODIMRLACYDTLV 106
Db 525 NHCGLSHEHIQNCRWDT-----EIKQLLTLSAMCONLARTAEDD--- 565
QY 107 HCETPAVTKRSIRLDETIWQTIKGPQVIYQETDPIF----- 146
Db 566 --EAP-----VLDNKLYLQIEKPYKEVMTRHVPVEELDSYHYQVELALQATENQHRV 615
QY 147 -----LMGNEKGMLTKKDAKOLEYAA-----KQFTPLSLSFDDLDRNNTPLWS 188
Db 616 DOVIKAVRKICSALDGVETPSVTEA-VKKLRAVNLPRNKSAADVTSLSGSDTRKNSTKGS 674
QY 189 SRPHNPMYVLP-----IFMHGKPNR-----SPNTPSHEAKOFTNEFRAPELKQV 234
Db 675 LNPNPQVQSDMHLTRIYDILLRLHANSRCSTGCPGRSRNIKEAWTATE-----QL 726
QY 235 SVKKAEDL---WCTSDLMFGYTOQSHWQIFNCKN-SRPF---RVHDYQPEIFLQ--- 285
Db 727 QFTVYAAHGISSNMVSNYEKYILCSLSH---NGKDLFKPIQSKKVGTYKNFFYLKWD 782
QY 286 ----PV-YSDLPWDGKYRMIGMGAVHSHNGESAKLSRSWN-----RAYL 324
Db 783 ELIIFPIQISQLPLESVLHLTLFGVLNQSNGSSPSDNKQKGPALGVSLTLDFKRF 842
QY 325 MAGMEWKNLTVMPRIWGRIFKEGSGQPDNDPIIDLYYGYGVDREPLYQLENKNSISGTVR 384
Db 843 TCG-----TKLLYLW-----TSSHTNSIPGAIPKKSVMYMERIVLQVDFPSPAFDIY 889
QY 385 YNPRSGKALQLDYVYPLGKIGSYFQIFQGYGQSLIDYNHEATSGV 432
Db 890 TSPQIDRNIIOODKLETLESIDK-----KLLDIHRDSSFGL 927

RESULT 2

US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170.996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0637000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-170-996-30

Query Match 4.3%; Score 101.5; DB 4; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.47;
Matches 98; Conservative 80; Mismatches 189; Indels 161; Gaps 21;

QY 1 MKVSL-----TLTSLSCFAILAIQQAQVNPVAFVD-----EYRSENDLQGD 46
Db 465 VKVSIIEGLQLPVFTTCDVSVSTEIIIMQALSWHDDLNQVDVGSYILKVCQGEVLQN 524
QY 47 NELPIDVQSATQSSTDPANPLDEHEPELYTTALENKTMLINCNSALNODIMRLACYDTLV 106
Db 525 NHCGLSHEHIQNCRWDT-----EIKQLLTLSAMCONLARTAEDD--- 565
QY 107 HCETPAVTKRSIRLDETIWQTIKGPQVIYQETDPIF----- 146
Db 566 --EAP-----VLDNKLYLQIEKPYKEVMTRHVPVEELDSYHYQVELALQATENQHRV 615
QY 147 -----LMGNEKGMLTKKDAKOLEYAA-----KQFTPLSLSFDDLDRNNTPLWS 188
Db 616 DOVIKAVRKICSALDGVETPSVTEA-VKKLRAVNLPRNKSAADVTSLSGSDTRKNSTKGS 674
QY 189 SRPHNPMYVLP-----IFMHGKPNR-----SPNTPSHEAKOFTNEFRAPELKQV 234
Db 675 LNPNPQVQSDMHLTRIYDILLRLHANSRCSTGCPGRSRNIKEAWTATE-----QL 726
QY 235 SVKKAEDL---WCTSDLMFGYTOQSHWQIFNCKN-SRPF---RVHDYQPEIFLQ--- 285
Db 727 QFTVYAAHGISSNMVSNYEKYILCSLSH---NGKDLFKPIQSKKVGTYKNFFYLKWD 782
QY 286 ----PV-YSDLPWDGKYRMIGMGAVHSHNGESAKLSRSWN-----RAYL 324
Db 783 ELIIFPIQISQLPLESVLHLTLFGVLNQSNGSSPSDNKQKGPALGVSLTLDFKRF 842
QY 325 MAGMEWKNLTVMPRIWGRIFKEGSGQPDNDPIIDLYYGYGVDREPLYQLENKNSISGTVR 384
Db 843 TCG-----TKLLYLW-----TSSHTNSIPGAIPKKSVMYMERIVLQVDFPSPAFDIY 889
QY 385 YNPRSGKALQLDYVYPLGKIGSYFQIFQGYGQSLIDYNHEATSGV 432
Db 890 TSPQIDRNIIOODKLETLESIDK-----KLLDIHRDSSFGL 927

RESULT 3

US-08-609-049A-13
; Sequence 13, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609.049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1658 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-049A-13

Query Match 4.3%; Score 100.5; DB 2; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.56;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTPLTSLSCFAILAIQ---QAQAVPNPVAFVDEYRS-----ENDLGQDNE 48
Db 381 LSPVTVQRMNCGENASVKVSIETGLQLPVTFCDVSSVTEIIMQALCWVHDDLQ--- 437

QY 49 LPIDVQS-----ATQSASTDTANPLDEHE--PELYTTALENKTMLNCSALNODINRLAC 101
Db 438 --VDVGSYILKVCQGEVLQNNHCLGSHEHIQNCRKWDTEIKLQLLTLSAMQNLARTAE 495

QY 102 YDTLVHGETPAVTKRSIRLDETWTQIKGKQVQIYQETDPF----- 146
Db 496 DD-----EAP-----VDLNKYLQIEKPYKEVMIRHPVEELDSYHYQVELALQTEN 542

QY 147 -----LMGNEKGMLTKDAKQLEYAA-----KQFTPLSLSFOLDLDRNN 183
Db 543 QHRAVDQVIKAVRKICSDALGDVETPSVTEA-VKKLRAVNLPRNKSAVTSLSGSDTRKN 601

QY 184 TPLWSSRPHPNPMVLP-----IFMHGKPNR-----SPNTPSHEAKQFTNPEFRAPE 229
Db 602 STKGSLNPNPQVQSMDLHTTAIYDLRLHLHANSRSCSTGCPGRSRNIKEAWTATE----- 656

QY 230 LKFOVSVKAAAEADL---WGTDSDLWFGYTOQSHWQIFNGKN-SRPF---RVHDIQPEIF 282
Db 657 ---QLQFTVYAAHGISSNNVSVYKYLICSLSH-----NGKDLFKPIQSKKVKTYKNFFY 709

QY 283 LTQ-----PV-YSDLPWDGKVRMIGMGAVHHNSGESAKLSRSWN----- 320
Db 710 LIKWDELIIFPIQISQLPLESVLHLLTFLGVNLQSSGSPDSNKKQKGPALGKVSILTLD 769

QY 321 -RAYLMAGMEWKNLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGDVRFYOLENKSNI 379
Db 770 FKRLFTCG-----TKLLYLW-----TSSHTNSIPGAIPKKSYYMERIVLQVDFPSPA 816

QY 380 SCTVRYNPRSGKALQOLDYVYPLGKIGISYFQIFQGYGOSLDYNEHATSFVG 432
Db 817 FDIYVTSPOIDRNIQODKLETLESIDK-----KLLDIHRDSSFGL 859

RESULT 4
US-09-170-996-13
; Sequence 13, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

```
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1658 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-170-996-13

Query Match 4.3%; Score 100.5; DB 4; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.56;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTPLTSLSCFAILAIQ---QAQAVPNPVAFVDEYRS-----ENDLGQDNE 48
Db 381 LSPVTVQRMNCGENASVKVSIETGLQLPVTFCDVSSVTEIIMQALCWVHDDLQ--- 437

QY 49 LPIDVQS-----ATQSASTDTANPLDEHE--PELYTTALENKTMLNCSALNODINRLAC 101
Db 438 --VDVGSYILKVCQGEVLQNNHCLGSHEHIQNCRKWDTEIKLQLLTLSAMQNLARTAE 495

QY 102 YDTLVHGETPAVTKRSIRLDETWTQIKGKQVQIYQETDPF----- 146
Db 496 DD-----EAP-----VDLNKYLQIEKPYKEVMIRHPVEELDSYHYQVELALQTEN 542

QY 147 -----LMGNEKGMLTKDAKQLEYAA-----KQFTPLSLSFOLDLDRNN 183
Db 543 QHRAVDQVIKAVRKICSDALGDVETPSVTEA-VKKLRAVNLPRNKSAVTSLSGSDTRKN 601

QY 184 TPLWSSRPHPNPMVLP-----IFMHGKPNR-----SPNTPSHEAKQFTNPEFRAPE 229
Db 602 STKGSLNPNPQVQSMDLHTTAIYDLRLHLHANSRSCSTGCPGRSRNIKEAWTATE----- 656

QY 230 LKFOVSVKAAAEADL---WGTDSDLWFGYTOQSHWQIFNGKN-SRPF---RVHDIQPEIF 282
Db 657 ---QLQFTVYAAHGISSNNVSVYKYLICSLSH-----NGKDLFKPIQSKKVKTYKNFFY 709

QY 283 LTQ-----PV-YSDLPWDGKVRMIGMGAVHHNSGESAKLSRSWN----- 320
Db 710 LIKWDELIIFPIQISQLPLESVLHLLTFLGVNLQSSGSPDSNKKQKGPALGKVSILTLD 769

QY 321 -RAYLMAGMEWKNLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGDVRFYOLENKSNI 379
Db 770 FKRLFTCG-----TKLLYLW-----TSSHTNSIPGAIPKKSYYMERIVLQVDFPSPA 816

QY 380 SCTVRYNPRSGKALQOLDYVYPLGKIGISYFQIFQGYGOSLDYNEHATSFVG 432
Db 817 FDIYVTSPOIDRNIQODKLETLESIDK-----KLLDIHRDSSFGL 859

RESULT 5
US-08-481-435-6
```

	Query Match	3.98;	Score 91.5;	DB 4;	Length 1024;
	Best Local Similarity	20.2%;	Pred. No. 2.2;		
	Matches	86;	Conservative	61;	Mismatches 172; Indels 107; Gaps 22;
Qy	1 MKVS-LSTLTLSLSCFAILLAIOQAQVAPNPVAFVDEVRSENLDQQDNELPI-DVQSATQ 58	: :			
Db	1 MKLSTTTTICLSISGAFGTTAI-----ALPTTVALLKNHQQNTERKQON--PIKDIFGLN 54	: :			
Qy	59 SASDSTDANPLDEHEPELYTTALENTMLINCSALNQDIRLCACYDTLVHGETTPAVIKTR 118	: : :			
Db	55 NVQVPNTIPL--HOTVEVT--NKAIYDYKDAQOKFEL-----AKSALNNKL 98	: : :			
Qy	119 SIRLDETITWOTKGKPQVIYQETTDPIFLMGNEKGMLTKDAKOLEYAANK-Q-FTPLSLUSF 177	: : :			
Db	99 QVEEDFKELLRT-----GVTNALNADLUKEWDIDTDLTFPNQSQFF 135	: : :			
Qy	178 DLDRNNTPLWSRRPHNPMYILP-IFMHGKPNRSP-NTPSHEAKOFTPFNEFRAPELKFOVS 235	: : :			

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:03 ; Search time 77.9 Seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLTLTLILSCFAILA.....YNHEATSFVGLMLNDWMGL 442

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_17.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	814	34.5	382	2	Q9K0U7	Q9K0U7 neisseria m
2	814	34.5	409	2	Q5UT21	Q9Jt21 neisseria m
3	372	15.8	306	2	Q9CL22	Q9CL22 pasteurella
4	367	15.6	329	2	Q9PMU8	Q9pmu8 campylobact
5	360.5	15.3	292	2	Q9SLD7	Q9slD7 versinia ps
6	360	15.3	289	2	Q9L6N9	Q9L6n9 salmonella
7	356.5	15.1	292	2	Q9Z4N8	Q9Z4n8 enterobacte
8	342.5	14.5	297	2	Q32349	Q32349 campylobact
9	246.5	10.4	355	2	Q25241	Q25241 helicobacte
10	240.5	10.2	355	2	Q9ZLX5	Q9ZLx5 helicobacte
11	122	5.2	278	2	Q9XB53	Q9xb53 erwiniia car
12	117	5.0	1686	4	O00443	O00443 homo sapien
13	112.5	4.8	602	11	Q63485	Q63485 rattus norv
14	110.5	4.7	821	2	Q59241	Q59241 bacillus sp
15	109	4.6	824	2	Q9F216	Q9f216 bacillus sp
16	107.5	4.6	901	5	O18749	O18749 caenorhabdi
17	105	4.4	435	2	Q9X9C0	Q9x9c0 streptococc
18	105	4.4	783	2	Q45554	Q45554 bacillus sp
19	103.5	4.4	1046	2	O84941	O84941 streptococc

20	103	4.4	798	4	Q9UN32	Q9un32 homo sapien
21	103	4.4	798	4	Q9UBK2	Q9ubk2 homo sapien
22	102.5	4.3	3247	12	O65553	O65553 bovine herp
23	101.5	4.3	403	5	P91736	P91736 hydra magni
24	101.5	4.3	660	2	Q9KGX8	Q9kgk8 bacillus ba
25	101.5	4.3	739	5	Q9V9E6	Q9v9e6 drosophila
26	101.5	4.3	838	10	O9C8I5	O9c8i5 arabidopsis
27	100.5	4.3	1509	11	O61194	O61194 mus musculu
28	100.5	4.3	1658	11	O61182	O61182 mus musculu
29	99.5	4.2	788	10	O9SCV5	O9scv5 arabidopsis
30	99	4.2	422	2	Q9RCZ6	Q9rcz6 streptomyce
31	99	4.2	5005	2	Q9PP25	Q9pp25 ureaplasma
32	98.5	4.2	467	4	Q9UJMS	Q9ujms homo sapien
33	98.5	4.2	467	4	Q9BTY2	Q9bty2 homo sapien
34	98.5	4.2	797	2	Q9RDW6	Q9rdw6 lactobacill
35	97.5	4.1	683	2	Q9A6U7	Q9a6u7 caulobacter
36	97	4.1	576	2	O9L1I5	O9l1i5 streptomyce
37	97	4.1	4307	5	Q9J3I9	Q9j3i9 caenorhabdi
38	96	4.1	871	4	Q9Y5C3	Q9y5c3 homo sapien
39	96	4.1	938	4	Q9Y5F7	Q9y5f7 homo sapien
40	95	4.0	690	5	O61142	O61142 plasmodium
41	95	4.0	719	2	O57114	O57114 streptococc
42	95	4.0	719	2	O9RET8	O9ret8 streptococc
43	95	4.0	791	10	O49137	O49137 arabidopsis
44	94.5	4.0	476	5	Q9BI60	Q9bi60 caenorhabdi
45	94.5	4.0	530	5	O45879	O45879 caenorhabdi

ALIGNMENTS

RESULT 1
Q9K0U7 PRELIMINARY; PRT; 382 AA.
AC Q9K0U7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."
RL Science 287:1809-1815(2000).
DR EMBL: AE002403; AAF40901.1; -.
DR TIGR: NMB0464; -.
DR InterPro: IPR003187; PLAL.
DR Pfam: PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 34.5%; Score 814; DB 2; Length 382;
Best Local Similarity 44.4%; Pred. No. 1e-62;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSALNDIMRLACYDYLHVGCTPAVI-----KTKRSIRLDETINQTI-KGKPOVIYQE 140
DB 33 LOCAALTDNVTIRACYDRIFAAQLPSSAGOGESKAVLNLTETVRSSLDKGEAVIVVEK 92

```

Qy 141 TTDPIFLMGNEKGMUTKKDAKOLEYAAKQFTPLSLSFDLDRNN-TPLWSSRPHNPMYVLP 199
Db 93 GGDAL-----PADSAGETADITYTPLSLMYDLDKDLRGLLGVREHNPMLMP 139

Qy 200 IFMHGKPNRSPNTPSH-EAKQFTPNFRAPELKFOVSVKVAEEDLWGTDSDLWFGYTOQ 258
Db 140 LWNNSPNYAPGSTRGTGTVQOEKFOQKRAETKLOVSEKSKIAEDLFKTRADLWFGYTQR 199

Qy 259 SHWQIFN-GKNSRPRRVHDYOPEIFLTQPVYSDLPWDGKVRMIGMGAVHHNGESAKLSR 317
Db 200 SDWQIYNGRKSAPPNTDYKPEIFLTQPVKADLPFGGRRLMGLGAGFVHQSNQGRPSR 259

Qy 318 SWNRAYLMAGMEWKNTVMPRIWGRIFKEGSGQPDNDPILDYGYGVDVRFYOLENKS 377
Db 260 SWNRAYMAGMEWGLTVIPRVVRAFDQ-SGDK-NDNPDADYMGYGDVKLQYRLNDRQ 317

Qy 378 NISGTVRNPNSGKALQOLDYVYPLGKIGISGYFOIQGYGQSLIDYNHEATSFVGVLMLN 437
Db 318 NVISVLRNPNTGYGAIEAAATFPIKGLKGVVRFHGYGESLIDYNHKNQNGIGIGLMFN 377

Qy 438 DWMGL 442
Db 378 DLDGI 382

RESULT 2
Q9JT21 PRELIMINARY; PRT; 409 AA.
AC Q9JT21
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Dellin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162757; CAB85240.1; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD585B064D01A41 CRC64;

```

```

Query Match 34.5%; Score 814; DB 2; Length 409;
Best Local Similarity 44.4%; Pred. No. 1.2e-62;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

Qy 87 INCSALNODIMRLACYDPLVHGETPAVI-----KTKRSIRLDETWTQTI-RGKQVYIQE 140
Db 60 LQCAALTDNVTRLACYDRIFAALPSSAGQEQESKAVNLTEIVRSSLDKGEAVIVVEK 119

Qy 141 TTDPIFLMGNEKGMUTKKDAKOLEYAAKQFTPLSLSFDLDRNN-TPLWSSRPHNPMYVLP 199
Db 120 GGDAL-----PADSAGETADITYTPLSLMYDLDKDLRGLLGVREHNPMLMP 166

Qy 200 IFMHGKPNRSPNTPSH-EAKQFTPNFRAPELKFOVSVKVAEEDLWGTDSDLWFGYTOQ 258
Db 167 LWNNSPNYAPGSTRGTGTVQOEKFOQKRAETKLOVSEKSKIAEDLFKTRADLWFGYTQR 226

```

```

Qy 259 SHWQIFN-GKNSRPRRVHDYOPEIFLTQPVYSDLPWDGKVRMIGMGAVHHNGESAKLSR 317
Db 227 SDWQIYNGRKSAPPNTDYKPEIFLTQPVKADLPFGGRRLMGLGAGFVHQSNQGRPSR 286

Qy 318 SWNRAYLMAGMEWKNTVMPRIWGRIFKEGSGQPDNDPILDYGYGVDVRFYOLENKS 377
Db 287 SWNRAYMAGMEWGLTVIPRVVRAFDQ-SGDK-NDNPDADYMGYGDVKLQYRLNDRQ 344

Qy 378 NISGTVRNPNSGKALQOLDYVYPLGKIGISGYFOIQGYGQSLIDYNHEATSFVGVLMLN 437
Db 345 NVISVLRNPNTGYGAIEAAATFPIKGLKGVVRFHGYGESLIDYNHKNQNGIGIGLMFN 404

Qy 438 DWMGL 442
Db 405 DLDGI 409

RESULT 3
Q9CL22 PRELIMINARY; PRT; 306 AA.
AC Q9CL22
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006179; AAK03510.1; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

```

```

Query Match 15.8%; Score 372; DB 2; Length 306;
Best Local Similarity 39.0%; Pred. No. 2.5e-24;
Matches 83; Conservative 40; Mismatches 82; Indels 8; Gaps 4;

Qy 229 ELKQVSVKVAEEDLWGTDSDLWFGYTOQSHWQIFNKGNSRPRRVHDYOPEIFLTQPVY 288
Db 97 EIKFKISLALPLWRGILGNNSVLAASYTKQSWFQSLNVDDSSPFRETNVEPQLFLAWKTQ 156

Qy 289 SLPWDGKVRMIGMGAVHHNG--ESAKLSRWNRAYLMAGMEWKNTVMPRIWGRIFKE 346
Db 157 YSLPFGWTLQDVETGINHQSNRDRDAEKLRSWNRLYVRAISAIKQNTVVEIKPMWRIPEK 216

Qy 347 GSGSQPDNDPILDYGYGVDVRF-LYOLENKSNTSGTVRNPNSGKALQOLDYVYPLGKG 405
Db 217 ---AKNDNDNPITYRGHFDVALGYHYHDHOFKLSG--HYNPISNKGGLSEASYPTKN 271

Qy 406 ISGYFOIQGYGQSLIDYNHEATSFVGVLMLND 438
Db 272 IRETVQYNGYGESLIDYQQRIGRIGIGISLNN 304

RESULT 4
Q9PMU8 PRELIMINARY; PRT; 329 AA.
AC Q9PMU8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

```

	PHOSPHOLIPASE A	(EC 3.1.1.32).	
DE	PIDA OR CJI351.		
GN	Campylobacter jejuni.		
OS	Bacteria; Proteobacteria;	epsilon subdivision; Campylobacter group;	
OC	Campylobacter.		
OX	NCBITaxID=197;		
RN	[]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCTC 11168;		
RX	MEDLINE=20150912; PubMed=10688204;		
RA	Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,		
RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,		
RA	Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,		
RA	Quail M.A., Rajadream M.A., Rutherford K.M., Van Vleet A.H.M.,		
RA	Whitehead S., Barrell B.G.;		
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni		
RT	reveals hypervariable sequences."		
RL	Nature 403:665-668(2000). ;		
DR	EMBL; ALI39078; CAB73778.1; .		
DR	InterPro; IPR003187; PLAI.		
DR	Pfam; PF02253; PLAI; 1.		
DR	PRINTS; PR01486; PHPLIPASEA1.		
KW	Complete proteome.		
SQ	SEQUENCE 329 AA; 38880 MW; A32731P2B751AC44 CRC64;		
Query Match 15.6%; Score 367; DB 2; Length 329;			
Best Local Similarity 33.5%; Pred. No. 7.5e-24;			
Matches 106; Conservative 47; Mismatches 125; Indels 38; Gaps			
QY	130 IKGKPVIVYTDPFLMGNEKGMLTKDAOLEYAAKQTPLSLSFOLDNRNT--PLM 187	:	: :
Db	43 LKNSSVLISEQNNSSOATQTQSITIKKEEK-----DFSRLALANYLGENESFNPLG 96	:	: :: :~
QY	188 SSRPNHMYVLPIFMHGCKNPSPNTTSHEAQTPTEFRAPELKFVOYSVKVAEADLWG 247	::	: :
Db	97 IS-SYKNNYF-P-FAYSFNSGLGVNNKSEA-----KFOLS VKRKLFE NLGL 141	:	: :~
QY	248 DSOLFYGTYQQSHWFNGFKNSPPRFVDHYDQPFIPTQVPY--SDLPWDGKVRMGMGAHV 306	:	: :~
Db	142 DEKYIAYTQSNWQIY--EHSSPFRETNYQPFFIDLPLYLKDYEFNNLR---VGILH 196	:	: :
QY	307 HSGESA-K--LSRSWNRAYLMAGWEKNLTVMPRINGRIKFEGSGSQPDNDPIDLDYYGY 364	:	: :
Db	197 ESVGKGDENLQRSWRNIYVSTAILYNKLFVPLRWIRI---PENKKDDDNPAILLHMGN 253	:	: :~
QY	365 GDVFRLFQLENKSNISGTVEYPNR--SGKCALODLYVYL-GKGISGYFOIQFGYCQLSI 421	:	: :
Db	254 FDVNLAY-LGGDTFYINLLRNLIKFNHNKGATQVDLGDFINFNGIYWYLOIFYNGYGESLI 312	:	: :
QY	422 DYNHATSFGVGLMLN 437		: :
Db	313 DYNKHQLRISTGF LIS 328	: :	: :
 RESULT 5			
ID	Q9SID7 PRELIMINARY; PRT; 292 AA.		
AC	Q9SID7;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DE	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PHOSPHOLIPASE A.		
GN	PIDA.		
OC	Yersinia pseudotuberculosis.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OX	Yersinia.		
OX	NCBITaxID=633;		
RP	[]		
RP	SEQUENCE FROM N.A.		
RC-	STRAIN=YPIII PIB1;		
RA	Karlyshev A.V., Winzeler E.A., Williams K.J., Oyston P.C.,		
RA	Titball R.W., Wren B.W.;		

RT	"Biochip-based Signature-Tagged Mutagenesis: Identification and									
RT	characterisation of Y. pseudotuberculosis gene plda essential for									
RT	virulence in mice.";									
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.									
DR	EMBL:	AJ245393;	CAB51586.1;	;						
DR	InterPro:	IPR003187;	PLAI:							
DR	Pfam:	PF02253;	PLAI:	1.						
DR	PRINTS:	PRO1486;	PHPLIPASEA1.							
SQ	SEQUENCE	292 AA;	33758 MW;	8E712D908ACB6BA5	CRC64;					
Qy	229	ELKFOVSVKVKAEDLNGTSDLWFGYQTQSHQIFNGKNSRPFVRVHDYQPEIFLTPV-	287							
Db	86	EVKFQLSLAFFPIWNGIAGDSLIGASTQKSWQAQSESSPFRETNYEPQLFLAWSTD	145							
Qy	288	YSDLPWGDGKVRMIGMAVHHSGESAKLSRSWNRAYLMAGMKNLTVMPRIWRIKKEG	347							
Db	146	YELAGW--TPREVEFGFNHQSNGKADPTSRSNRVYTRVMAQRGNLEIDLKWPYRIPESD	203							
Qy	348	SGSQPDNDPILDYGYGDVRFYQL--ENKSNTSGTVRYNPRSGKALQLDYYVPLGKI	406							
Db	204	S---KDDNPDTIKMGYRLKVGVALGDSVFSLDG--RYNMNTGYGGAENGWSYPTIKHV	258							
Qy	407	SGYFIQFGYGGQSLIDYNHEATSGVGLMLND	438							
Db	259	RFYQVFSGGYSGESMIDYFNRQTRVGVGIMLND	290							
RESULT	6									
Q9L6N9		PRELIMINARY;	PRT;	289	AA.					
ID	Q9L6N9									
AC	Q9L6N9;									
DT	01-OCT-2000 (TrEMBLrel. 15, Created)									
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)									
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)									
DE	PLDA PROTEIN.									
GN	PLDA.									
OS	Salmonella typhimurium LT2.									
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;									
OC	Salmonella.									
OX	NCBI_TaxID=99287;									
ON	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=SGSC1412;									
RA	Washu;									
RL	"The Salmonella typhimurium Genome Sequencing Project.";									
RT	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=SGSC1412;									
RA	Waterston R.;									
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.									
DR	EMBL:	AF233324;	AAF33435.1;	;						
DR	InterPro:	IPR003187;	PLAI:							
DR	Pfam:	PF02253;	PLAI:	1.						
SQ	SEQUENCE	289 AA;	32967 MW;	DA97F5E1651C49C6	CRC64;					
Qy	226	RAPELKFOVSVKVKAEDLNGTSDLWFGYQTQSHQIFNGKNSRPFVRVHDYQPEIFLTQ	285							
Db	83	RKDEVKFQLSLAFFPIWNGIAGDSLIGASTQKSWQAQSESSPFRETNYEPQLFLG	142							
Qy	286	PV-YSDLPWGDGKVRMIGMAVHHSGESAKLSRSWNRAYLMAGMKNLTVMPRIWRIK	344							
Db	143	ATDYRFAGW--TLRDVEMGYNHDSGRDPTSRSNRNLRYRLMAENGWLVEVKPWVI-	199							
Qy	287	YSDLPWGDGKVRMIGMAVHHSGESAKLSRSWNRAYLMAGMKNLTVMPRIWRIK	344							
Db	144	ATDYRFAGW--TLRDVEMGYNHDSGRDPTSRSNRNLRYRLMAENGWLVEVKPWVI-	199							

```
Qy 345 KEGSGQDDNPDIIDLYYGYGDRVLYOLENKSNIQSVTVRNPRSGKALOLDYVYPLGK 404
Db 200 -----GSTDDNPDIITYKGYQLKGYHL-GEAVLSAKQYNWNTGYGAEVGLSYPTK 253

Qy 405 GISGYFQIFQGYGOSLIDYNHEATSFVGLMLND 438
Db 254 HVRLYTQVSYGSGESLIDYNFQTRVGVGLMLND 287

RESULT 7
Q924N8 PRELIMINARY; PRT; 292 AA.
AC Q924N8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (EC 3.1.1.32).
GN PLDA.
OS Enterobacter agglomerans.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G., Brinkman E., van Boxel R., Bekkers A.C., Verheij H.M.,
RA Tommassen J.;
RT "Molecular characterization of enterobacterial pida genes encoding
RT outer membrane phospholipase A.,"
RL J. Bacteriol. 176:861-870(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120145; PubMed=9921577;
RA Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tommassen J.;
RT "Sequence comparison of outer membrane phospholipases A: implications
RT for structure and for the catalytic mechanism.,"
RL Res. Microbiol. 149:703-710(1998).
DR EMBL; AF034414; AAD03498.1; -.
DR InterPro; IPR003187; PLAI.
DR Pfam; PF02253; PLAI.
DR PRINTS; PR01486; PHPLIPASEAL.
KW Signal; Hydrolase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 292 OUTER MEMBRANE PHOSPHOLIPASE A.
SQ SEQUENCE 292 AA; 33719 MW; B75516D093B2BEEA CRC64;

Query Match 15.18; Score 356.5; DB 2; Length 292;
Best Local Similarity 40.18; Pred. No. 5.2e-23;
Matches 85; Conservative 31; Mismatches 87; Indels 9; Gaps 5;

Qy 229 ELKFQSVKVAEDLWCTGDSLWFGYQQSHWQIFNCKNSRPRFRVHDYQPEIFLTQPV- 287
Db 86 EVKQLSLGFFPIWRCIAGNSLLGASYTQRSGWQANSDESSPFRETNYEPOIFLAWTD 145

Qy 288 YSCLPQWCKVRMIGMGAHVHNGESAKLSRSWNRAYLMAGMEWKNLTVMYPRTWGRIFKEG 347
Db 146 YELAGW--TFREVEFGYNHQSNRGPDSRSMDRYTTRLMAQRGNLEIDLPWYRIPESD 203

Qy 348 SGSPQDDNPDIIDLYYGYGDRVLYOLENKSNIQSVTVRNPRSGKALOLDYVYPLGKI 406
Db 204 S---KDDNPDIINKYGYRLRVGVALGESVFSIDG--RYNNTGYGGAEMGWSYPTKHV 258

Qy 407 SGYFQIFQGYGOSLIDYNHEATSFVGLMLND 438
Db 259 RFTYQVFGSGESMIDYNFQTRVGVGLMLND 290

RESULT 8
O32349 PRELIMINARY; PRT; 297 AA.
ID O32349
AC O32349;
```

```
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Campylobacter coli.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=195;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=UA585;
RX MEDLINE=97230284; PubMed=9119448;
RA Grant K.A., Belandia I., Dekker N., Richardson P.T., Park S.F.;
RT "Molecular characterization of pida, the structural gene for a
RT phospholipase A from Campylobacter coli, and its contribution to cell-
RL associated hemolysis.,"
RL Infect. Immun. 65:1172-1180(1997).
DR EMBL; Y11031; CAA71915.1; -.
DR InterPro; IPR003187; PLAI.
DR Pfam; PF02253; PLAI.
DR PRINTS; PR01486; PHPLIPASEAL.
SQ SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;

Query Match 14.58; Score 342.5; DB 2; Length 297;
Best Local Similarity 32.48; Pred. No. 8.8e-22;
Matches 97; Conservative 37; Mismatches 108; Indels 57; Gaps 11;

Qy 178 DLDRNPTPLWSRRPHNMYVLPIFMHGKPNRSPNT-----P 213
Db 16 DLKENNASLLSRKHETQ-----NTQKTPSTKEDFSRIALANYLGENSESFNPLGIS 66

Qy 214 SHEAKQFTP-----NEFRAPELKFOVSVKVKAEDLWCTGDSLWFGYQQSHWQIF 264
Db 67 SYKNMYFLPFAYSGSLGGENRKTEMKFLSIKKRFLFDLLGLGKGYVGYTQTSWQ-- 124

Qy 265 NGKNSRPRFRVHDYQPEIFLTQPV-YSDLPWCKVRMIGMGAHVHNGESAK--LSRSWNR 321
Db 125 NYKHSSPFRETNYQPEFFVDIPLHEDYKFLNLR--VGILHESNGKGDENLESRSNR 181

Qy 322 AYLMAWMEWKNLTVMYPRTWGRIFKEGSGQDDNPDIIDLYYGYGDRVLYOLENKSNIQSG 381
Db 182 IYASVFLYQRFLEFVPRYIYRI---PENSEDDNDNPETHYMGNFIDIN-IGSLGNDYFINL 237

Qy 382 TVRYNP--RSCKGKALQLDYVYPL-CKGISGYFQIFQGYGOSLIDYNHEATSFVGLMLN 437
Db 238 MERNLDFHDKNGAVQVDIGDIFDNGIYWLQYFNGYGGSLIDINKRLQLRSLTAFLIS 296

RESULT 9
O25241 PRELIMINARY; PRT; 355 AA.
AC O25241;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).
GN HP0499.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
```


RA	Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA	Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA	Venter J.C.;
RT	"The complete genome sequence of the gastric pathogen <i>Helicobacter</i>
RT	<i>pylori</i> .";
RL	Nature 388:539-547(1997).
DR	ENBL: AE000564; AAD07564.1; -.
DR	TIGR: HP0439; -.
DR	InterPro: IPR003187; PLAI.
DR	Fram: PF02253; PLAI; 1.
KW	Hypothesized protein; Complete proteome.
SW	SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match	10.4%	Score 245.5;	DB 2;	Length 355;
Best Local Similarity	25.1%;	Pred. NO. 2.6e-13;		
Matches 82;	Conservative 41;	Mismatches 117;	Indels 87;	Gaps 11;

[illegible]

RESULT	10		
Q9ZLX5			
ID	Q9ZLX5	PRELIMINARY;	PRT; 355 AA.
AC	Q9ZLX5;		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. -10, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE	PUTATIVE PHOSPHOLIPASE A1.		
GN	PLDA OR JHP0451.		
OS	Helicobacter pylori J99 (Campylobacter pylori J99).		
OC	Bacteria; Proteobacteria;	epsilon subdivision; Helicobacter group;	
OC	Helicobacter.		
OX	NCBI TaxID=85963;		

RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; Pubmed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uilla-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voris G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen *Helicobacter pylori*.";
RL Nature 397:176-180(1999).
DR ENBL: AE001479; AAD06029.1; -.
DR InterPro: IPR003187; PLAL.
DR Pfam: PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRG64;

Query Match	10.2%	Score 240.5;	DB 2;	Length 355;
Best Local Similarity	25.1%;	Pred. NO. 8.8e-13;		
Matches 83;	Conservative 44;	Mismatches 109;	Indels 95;	Gaps 13;

[illegible]

RESULT	11	
Q9XB53	PRELIMINARY;	PRT; 278 AA.
Q9XB53		
AC	Q9XB53;	
DT	01-NOV-1999	(TREMBLrel. 12, Created)
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)
DT	01-NOV-1999	(TREMBLrel. 12, Last annotation update)
DE	KDUI.	
GN	KDUI.	
OS	Erwinia carotovora.	
OC	Pectobacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;	
OC	Bacterium.	
OX	NCBI_TaxID=554;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 39048, GS101;	
RX	MEDLINE=98065591; PubMed=9402024;	
RA	McGowan S.J., Sebalhia M., O'Leary S., Hardie K.R., Williams P.,	
RA	Stewart G.S., Bycroft B.W., Salmond G.P.;	
RT	"Analysis of the carbenem gene cluster of Erwinia carotovora:	
RT	definition of the antibiotic biosynthetic genes and evidence for a	
RT	novel beta-lactam resistance mechanism.";	
RL	Mol. Microbiol. 26:545-556(1997).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 39048, GS101;	
RX	MEDLINE=98276484; PubMed=9614345;	
RA	McGowan S.J., Bycroft B.W., Salmond G.P.;	
RT	"Bacterial production of carbenams and clavams: evolution of beta-	
RT	lactam antibiotic pathways.";	
RL	Trends Microbiol. 6:203-208(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 39048, GS101;	
RA	McGowan S.J.;	
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; U17224; AAD38237.1; -.	
SQ	SEQUENCE 278 AA; 31341 MW; 63769FAF3550E1B7 CRC64;	

Query Match 5.2%; Score 122; DB 2; Length 278;
Best Local Similarity 21.3%; Pred. No. 0.013;
Matches 73; Conservative 53; Mismatches 128; Indels 88; Gaps 16;

DR PROSITE; PSS0290; PI3_4_KINASE_3; 1.
SQ SEQUENCE 1686 AA; 190736 MW; E9311C803025C96F CRC64;

Query Match 4.8%; Score 112.5; DB 11; Length 602;
Best Local Similarity 19.3%; Pred. No. 0.27;
Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;

6	TQOLISSILILVL-LLSLPTALAAEGNTREDNFKLLGNDNVKVRSEAGALQLEQVDGQ	64
66	NPL-DEHEPELY-----TTALENKTMLINGSAL-----NODIMKLACY-DTLVHGTEP	111
65	MTLVQDHQGEKIQLRGSMTHGLQWFPETLNDNAYKALANDWESNMIRLAMYVGENGYSNP	124
112	AVIKTKRSIRLDETI-----WQT-----IKGKQPVIV	138
125	ELTKSRVIGKIDIAIENDMTYVVDVHVHAPGDRPDVPYAGADEFRDIAALYNPNPHIY	184
139	QETTPD-----IFLMGNEKGMLTKKDAKQLEYAAKQFTPLSLSPDLDRN-----NTPLW	187
185	ELANEPSSNNGGAGIPNNEEGNAVK-----EYADPIVEMLRDSGNADDNIIIVGSPNW	239
188	SSRP-----HNPWYVLPF--MHGKPNRS--PNTPSHAKQFTNPEFPAPELKFQ	233
240	SQRDLAADNPIDDHTTMYTVHYTGSAASTESYPETPNSEGNVMSNTRYA-----LE	295
234	VSVKVAEAEDLWGT-----DSDLWFQYTOQ-----SHWQIFNGKNS-----RP	271
296	NGVAVFATE--WCTSQANGDGGPYFDEADVWIEFLNENNISWANWSLTN-KNEVSGAFTP	352
272	FRVH-----DYQPEIFLTQPYSDLPW---D	294
353	FELGKSNAATSLDPGDPQVWVPEELSLSGEYVRARIKGVNYEP---IDRTKYTKVLDFND	409
295	GKVRMIGMGVAHHSGRS-----AKLSRSNRWAYLMAGMEKNLNTVMPRIWG	341
410	GTKQCGFV-----NGDSPVEDVVIENEAGALKSLGLDASNDVSEGNVHANARLSADGWG	463
342	R 342	
464	K 464	

RESULT 15
O9F216

AD	Q9F216;	
AC	PROLIMINAL;	FRI; 824 AA.
DT	01-MAR-2001 (Tremblrel. 16, Created)	
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)	
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)	
DE	CELLULOSE.	
DE	Bacillus sp.	
OS	Bacteria; Firmicutes; Bacillus/Clostridium group;	
OC	Bacillus/Staphylococcus group; Bacillus.	
OX	NCBI_TaxID=1409;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=KSM-5237;	
RX	MEDLINE=21036886; PubMed=11193393;	
RA	Hakamada Y., Hatada Y., Kolke K., Yoshimatsu T., Kawai S.,	
RA	Kobayashi T., Ito S.;	
RT	"Reduced amino acid sequence and possible catalytic residues of a	
RT	thermostable, alkaline cellulase from an alkaliphilic Bacillus	
RT	strain.";	
RL	Biosci. Biotechnol. Biochem. 64:2281-2289(2000).	
RL	EMBL=AB018420; BAB19360.1; -	
DR	InterPro: IPR001547; Glyco_hydro_F5.	
DR	Pfam: PF00150; cellulase; 1.	
DR	PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.	
DR	SEQUENCE 824 AA; 91564 MW; 65FA940FE1D729B9 CRC64;	

```

Query Match      4.68; Score 109; DB 2; Length 824;
Best Local Similarity 20.08; Pred. No. 0.85;
Matches 96; Conservative 59; Mismatches 159; Indels 166; Gaps 28;

Qy 9 TLTSLTSCFATLAIQQAAQVNPVAFVDEVRSEND--LGDQNELPIDVOSATQSASTDTA 65
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 7 TKQLSTLLIIVL-LLSLFLPAALAEAGNTRDEGNFKHLILGNNDVNRKPSAGALQLQEVGQ 65

Qy 66 NPL-DEHEPELY-----TTALENKTMLINC SAL-----NODIMRLACYDTLVHGE--- 109

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:53 ; Search time 28.02 Seconds
(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 2360

Sequence: 1 MKVSLSTLTSLILSCFAILA.....YNHEATSGVGLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	373	15.8	286	PAL_KLEPN	P37446 klebsiella
2	358	15.2	289	PAL_SALTY	P37442 salmonella
3	357	15.1	289	PAL_ECOLI	P00631 escherichia
4	321.5	13.6	289	PAL_PROVU	P37447 proteus vul
5	110.5	4.7	800	GUN_BACS1	P06564 bacillus sp
6	98.5	4.2	810	ANDM_YEAST	P15274 saccharomyc
7	96	4.1	888	YJH0_YEAST	P40361 saccharomyc
8	94.5	4.0	397	GUN_PAEPO	P23548 paenibacilli
9	93.5	4.0	1478	BCK1_YEAST	Q01389 saccharomyc
10	93	3.9	716	BAC2_MOUSE	P97303 mus musculu
11	92.5	3.9	849	SRK6_BRAOL	Q09092 brassica ol
12	92	3.9	324	Y170_ARCFU	O28112 archaeoglob
13	92	3.9	1158	RI14_HUMAN	P48552 homo sapien
14	92	3.9	1788	YP72_CAEEL	Q09221 caenorhabdi
15	91.5	3.9	467	INVO_MOUSE	P48997 mus musculu
16	91.5	3.9	486	VG14_BPM15	Q05220 mycobacteri
17	91.5	3.9	1024	Y075_MYCGE	P47321 mycoplasma
18	90.5	3.8	379	YJHT_HAEIN	P44544 haemophilus
19	90	3.8	496	CATA_DICDI	O77229 dictyosteli
20	90	3.8	719	PBPA_STRPN	Q04707 streptococc
21	89.5	3.8	503	CD44_RAT	P26051 rattus norv
22	89.5	3.8	825	GUN3_BACS4	P19570 bacillus sp
23	89	3.8	521	NPRE_BACAM	P06832 bacillus am
24	88.5	3.8	1379	MET_MOUSE	P16056 mus musculu
25	88	3.7	1131	PMAL_DUNBI	P54211 dunaliella
26	87.5	3.7	353	DCUP_BACSU	P32395 bacillus su
27	87.5	3.7	478	DHGB_ACICA	P13650 acinetobact
28	87.5	3.7	669	AMY_ALTHA	P29957 alteromonas
29	87.5	3.7	794	SEIL_HUMAN	Q9ubv2 homo sapien
30	87.5	3.7	969	SACB_STRSL	Q55242 streptococ
31	87	3.7	483	MURE_CHLTR	O84271 chlamydia t
32	87	3.7	537	P4H2_MOUSE	Q60716 mus musculu
33	87	3.7	560	DTXH_CORBE	P00589 corynebphage

Query Match 15.8%; Score 373; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 8.1e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

ALIGNMENTS

RESULT 1

ID	PAL_KLEPN	STANDARD;	PRT;	286 AA.
AC	P37446;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PLDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
OX	NCBI_TaxID=573;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94131966; PubMed=8300539;			
RA	Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.;			
RT	"Molecular characterization of enterobacterial plda genes encoding			
RT	outer membrane phospholipase A.;"			
RL	J. Bacteriol. 176:861-870(1994).			
CC	- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- COPACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X76901; CAAS4223.1; ..			
DR	PIR; B36971; B36971.			
DR	PIR; S40129; S40129.			
DR	InterPro; IPR003187; PLAL.			
DR	Pfam; PF02253; PLAL; 1.			
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.			
FT	SIGNAL 1 20			
FT	CHAIN 21 286			
FT	ACT_SITE 161 161			
FT	BY SIMILARITY.			
FT	PHOSPHOLIPASE A1.			
SEQUENCE	286 AA; 32544 MW; 3E39F863085108A3 CRC64;			

```
QY 226 RAPELKFOVSVKAAEDLWGTDSLWFGYTOQSHWQIFNGKNSRPFVRVHDYQEIFLTQ 285
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
Db 80 RKDEVKFLSLAFPLWRGLGDSILGASYTKQSKWQLSNSKESAPFRETNYPOLPLGF 139
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
QY 286 PV-YSDLPWDKVRMIGMAVHSHNGESAKLSRSNNRAYLMAGMEWKNTVMRPIWGRIF 344
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
Db 140 ATDYQFAGW--TLRDIEMGYNHDSNGRSDPTSRWNRLYARLMAQNGNWLVEVKPW--Y 194
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
QY 345 KEGSSQDDNPDLIDYGYGDVRFYOLENKSNTSGTVNPRSGKQALQLDYVYPGLK 404
    || ||||| ||| : : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db 195 VVGs---TDDNPDIYKMGYTRKVGYLQ-GEALUSAGQYNWNTGYGGAELGVSPITK 250
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
QY 405 GISGYFQIFOGYGQSLDYNHEATSFVGGLMND 438
    : | : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db 251 HVRAYTQIYSGYGESLIDYNFNQTRVGVGLMND 284
    : | : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||

RESULT 2
PAL_SALTY
ID PAL_SALTY STANDARD; PRT; 289 AA.
AC P37442;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tomassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COPACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X76900; GAA54222.1; -.
DR PIR; A36971; A36971.
DR PIR; S40131; S40131.
DR StyGene; SG10302; plda.
DR InterPro; IPR003187; PLA1.
DR Pfam; PF02253; PLA1; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1
FT CHAIN 20
FT ACT_SITE 164 164 BY SIMILARITY.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 289 AA; 33003 MW; D5ACF5E33F704DBA CRC64;
```

Query Match

15.2%; Score 358; DB 1; Length 289;

```
Best Local Similarity 39.3%; Pred. No. 1.3e-21;
Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;
QY 226 RAPELKFOVSVKAAEDLWGTDSLWFGYTOQSHWQIFNGKNSRPFVRVHDYQEIFLTQ 285
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
Db 83 RKDEVKFLSLAFPLWRGLGILGNPNSILGASYTKQSKWQLSNSKESPPFRETNYPOLPLGF 142
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
QY 286 PV-YSDLPWDKVRMIGMAVHSHNGESAKLSRSNNRAYLMAGMEWKNTVMRPIWGRIF 344
    | : ||| : : : | : ||| : : | : ||| : : | : ||| : : | : ||| : : |
Db 143 ATDYQFAGW--TLRHEVEMGYNHDSNGRSDPTSRWNRLYARLMAQNGNWLVEVKPWYI- 199
    || ||||| ||| : : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
QY 345 KEGSSQDDNPDLIDYGYGDVRFYOLENKSNTSGTVNPRSGKQALQLDYVYPGLK 404
    ||||| ||| : : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db 200 ----GSTDNDPDIYKMGYQYQKIGYHL-GEAVLSAKQYNWNTGYGGAELGVSPVTK 253
    : | : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
QY 405 GISGYFQIFOGYGQSLDYNHEATSFVGGLMND 438
    : | : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db 254 HVRLYQIYSGYGESLIDYNFNQTRVGVGLMND 287
    : | : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||

RESULT 3
PAL_ECOLI
ID PAL_ECOLI STANDARD; PRT; 289 AA.
AC P00631;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA OR B3821 OR Z5342 OR ECS4751.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85157492; PubMed=6397464;
RA Honma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,
RA Inoue K., Ikeda H., Sekiguchi M., Nojima S.;
RT "The DNA sequence encoding plda gene, the structural gene for
RT detergent-resistant phospholipase A of E. coli.";
RL J. Biochem. 96:1655-1664(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes.";
RL Science 257:771-778(1992).
RN [3]
RP REVISION TO 14-15.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
```

RL Nature 409:529-533(2001).
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RA Hayashi T., Makino K., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 30-289 FROM N.A.
RC STRAIN=K12;
RA MEDLINE=85003590; PubMed=6383820;
RX de Geus P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
RT "The pro- and mature forms of the E. coli K-12 outer membrane
RT phospholipase A are identical.";
RL EMBO J. 3:1799-1802(1984).
RN [7]
RP SEQUENCE OF 174-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87115164; PubMed=3027506;
RA Irino N., Nakayama K., Nakayama H.;
RT "The recQ gene of Escherichia coli K12: primary structure and
RT evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).
RN [8]
RP MUTAGENESIS OF SER-172.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RT Verheij H.M., Tommassen J.;
RA "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A";
RL J. Bacteriol. 176:861-870(1994).
RN [9]
RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
RX MEDLINE=91249806; PubMed=2040286;
RA Horrevorts A.J.G., Verheij H.M., de Haas G.H.;
RT "Inactivation of Escherichia coli outer-membrane phospholipase A by
RT the affinity label hexadecanesulfonyl fluoride. Evidence for an
RT active-site serine.";
RL Eur. J. Biochem. 198:247-253(1991).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
CC DORMANT IN NORMAL GROWING CELLS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
CC LYSIS OR TEMPERATURE SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02143; CAA26081.1; -
DR EMBL; M87049; AAA67617.1; -
DR EMBL; AE000458; AAC76824.1; -
DR EMBL; AE005613; AAG59017.1; -
DR EMBL; AP02567; BAB38174.1; -
DR EMBL; M30198; AAA24516.1; -
DR PIR; A00771; PSECA.

DR PIR; A22133; PSECAL.
DR PIR; S30711; S30711.
DR EcoGene; EGI0738; plga.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
FT Complete proteome.
FT SIGNAL 1 20 PHOSPHOLIPASE A1.
FT CHAIN 21 289
FT ACT_SITE 164 164
FT MUTAGEN 172 172
FT CONFLICT 14 15 S->F: INACTIVE PROTEIN.
FT CONFLICT 30 33 LP -> FA (IN REF. 2).
FT SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CRC64;
Query Match 15.1%; Score 357; DB 1; Length 289;
Best Local Similarity 39.3%; Pred. No. 1.6e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;
QY 226 RAPELKFOVSKVKAADLWGTDSLWFGYTCQSHWQIENGKNSRPRFRVHDYQPEIFLTQ 285
DB 83 RKDEVKFOISLAFPLMRGILGPSNLGASVYQKSWQLSNSESSPFRNYPQLGLGF 142
QY 286 PV-YSDLPMDGKVRMIGMGAVHHSNGESAKLSRSWNRAYLMAGMEKKNLTVMPRIINGRIF 344
DB 143 ATDYREFAGW--TLRDVEMGYNHDSNGRSDPTSRWNRLYTRLMAENGWLVKPYVW- 199
QY 345 KEGSGQPDNDPILDYGYGVDFRFLYQLENKSNISCTVYRNPNSGKALQLODYVPLGK 404
DB 200 -----GNTDDNDPDTTMYGYKIGYHL-GDAVLAKGQYNWNTGYGGAELGLSPDTK 253
QY 405 GISGYQIFQGYGQSLIDYNHEATSFQVGLMLND 438
DB 254 HVRLYQVSYGYESLIDYNFNQTRVGVGVMLND 287
RESULT 4
PAL_PROVU STANDARD; PRT; 289 AA.
AC P37447;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A";
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02143; CAA26081.1; -
DR EMBL; M87049; AAA67617.1; -
DR EMBL; AE000458; AAC76824.1; -
DR EMBL; AE005613; AAG59017.1; -
DR EMBL; AP02567; BAB38174.1; -
DR EMBL; M30198; AAA24516.1; -
DR PIR; A00771; PSECA.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X76902; CAA54224.1; -;
 DR PIR; C36971; C36971.
 DR PIR; S40130; S40130.
 DR InterPro; IPR003187; PLAL.
 DR Pfam; PF02253; PLAL; 1.
 KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
 FT SIGNAL 20 BY SIMILARITY.
 FT CHAIN 21 289 PHOSPHOLIPASE A1.
 FT ACT_SITE 164 164 BY SIMILARITY.
 SQ SEQUENCE 289 AA; 32944 MW; D75516CFFB406997 CRC64;

Query Match 13.6%; Score 321.5; DB 1; Length 289;
 Best Local Similarity 33.3%; Pred. No. 1.le-18;
 Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

QY 179 LDRNNTPLWSSRPMPVPLTFMFHGKPNRSPNTPSHEAKQFTPNFRAPELKFQVSVKV 238
 DB 43 LQEHNP-FTLYPYSNLLIY-----TSDLNKAIESYNSDNA-NKDEVKFLSLAF 95
 QY 239 KAAEDLWGTDSLWFGYTOQSHQWIFNGKNSRPFVRVDYQPEIFLTQPV-YSDLPWQGV 297
 DB 96 PLWRGILGDSNLGASYTQSRWSQWLSNTGESAPFRETNYEPQLFGLFATDYSVGDW--TL 153
 QY 298 RMIGMGAVHNSGESAKLSRSNRAYLMAGMEWKNLTVMPTWGRIFKEGSGQDDNDPD 357
 DB 154 RDAEFGYHQSNRSDPTSRNRLYSRLMAQNGNWLVEVKPVI-----GDTSDNKN 207
 QY 358 IDYYGYGDVRLYLENKSNTSGVRYNPRSGKALQLDYVPLGKISGYFQIFQGVG 417
 DB 208 ITKMGYQKIGYQL-GEVLASAGQYNWNTGYGAELGVSYPITKHFRTYQVSGTG 266
 QY 418 QSLIDYNHEATSFYGLMLND 438
 DB 267 ESLIDYDFNOTRVGVMVND 287

RESULT 5
 GUN_BACSL STANDARD; PRT; 800 AA.
 ID GUN_BACSL
 AC P06564;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
 DE (ALKALINE CELLULOSE).
 OS Bacillus sp. (strain 1139).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1411;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87085443; PubMed=3098909;
 RA Fukumori F., Kudo T., Narahashi Y., Horikoshi K.;
 RT "Molecular cloning and nucleotide sequence of the alkaline cellulase
 gene from the alkalophilic Bacillus sp. strain 1139.";
 RL J. Gen. Microbiol. 132:2329-2335(1986).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
 LINKAGES IN CELLULOSE.
 CC -1- MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES
 ONE ALKALINE CELLULOSE (PH OPTIMUM 9.0). THIS STRAIN IS NOT A TRUE
 CELLULOTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO
 HYDROLYSE NATIVE CELLULOSE.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
 HYDROLASES).
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M15743; AAA22305.1; -;
 DR EMBL; D00066; BAA00045.1; -;
 DR PIR; A29003; A29003.
 DR InterPro; IPR001547; Glyco_hydro_F5.
 DR Pfam; PF0150; cellulase; 1.
 DR PROSITE; PS00659; GLYCOSYL-HYDROL_F5; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 800 ENDOGLUCANASE.
 FT ACT_SITE 190 190 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 305 305 NUCLEOPHILE (BY SIMILARITY).
 SQ SEQUENCE 800 AA; 88602 MW; 7CCA4D7B6DAD55CF CRC64;

Query Match 4.7%; Score 110.5; DB 1; Length 800;
 Best Local Similarity 20.0%; Pred. No. 0.42;
 Matches 96; Conservative 53; Mismatches 163; Indels 169; Gaps 25;

QY 9 TSLISCFALIAOQAQVNPVAFVDEYRSEN-----LGQDNELPIDVQSASTDTA 65
 DB 7 TKQLISSILILVL-LLSLFPPTALAAEGNTREDNFKHLGNDNVKRPSEAGALQLQEVQ 65
 QY 66 NPL-DEHEPELY-----TTALENKTMLNCSAL-----NQDINRLACY-DTLVHGETP 111
 DB 66 MTLVDQHGKEIKQLRGMTGHLQWFFELINDNAYKALANDWESNMIRLAMYGENGYASN 125
 QY 112 AVIKTKRSIRLDETI-----WQT-----IKGKPOVIY 138
 DB 126 ELIKSRVIGKIDLAENDMYVIVDWHVHAPDRDPVYAGAEFFRDIAALPNPHIY 185
 QY 139 QETDP-----IFLMGNEKGLTKDKAQLEYAAKQFTPLSLSFDLDRN-----NTPW 187
 DB 186 ELANEPSSNNNGGAGIPNNEGNAVK-----EYADPIVEMLRDSGNADDNIIIVGSPNW 240
 QY 188 SSRP-----HNPXVLPPIF--MHGKPNRS--PNTPSHEAKQFTPNFRAPELKFQ 233
 DB 241 SQRPDLAADNPIDHHHTYTVHFYTGSHAASTESYPPETPNSENGVMSNTRYA-----LE 296
 QY 234 VSVKVAEADLWGT-----DSDLWFGYTOQ-----SHWQIFNGKNS-----RP 271
 DB 297 NCVAVFATE--WGTQANGCGPYDEADVWIEFLNENNISWANNSLTN-KNEVSCATP 353
 QY 272 FRVH-----DYOPEIFLTQPVYSDLPW---D 294
 DB 354 FELGKSNATSLDPGPDQVWVPELSLSGEYVRARIKGVNYEP---IDRTKYTKVLWDFND 410
 QY 295 GKVRMIGMGAVHNSGES-----AKLSRSNRAYLMAGMEWKNLTVMPTW 341
 DB 411 GTKQGFV-----NGDSPVEDVWIENEAGALKSLDASNDVSEGNVWANRLSADGM 464
 QY 342 R 342
 DB 465 K 465

RESULT 6

AMDM_YEAST STANDARD; PRT; 810 AA.
 ID AMDM_YEAST
 AC P15274;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE AMP DEAMINASE (EC 3.5.4.6) (MYOADENYLATE DEAMINASE).
 GN AMD1 OR AMD OR YKL035C
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

P97303;
15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2).
BACH2.
GN GN
MS musculus (Mouse).
OS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N. A.
RP
RC STRAIN=BALB/C;
RX MEDLINE=97042438; PubMed=8807638;
RY Oyake T., Itoh K., Mochizashi H., Hayashi N., Hoshino H., Nishizawa M.,
RA Yamamoto M., Igarashi K.;
RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
transcription factors that interact with Mafk and regulate
transcription through the NF-E2 site.";
RL Mol. Cell. Biol. 16:6083-6095(1996).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
CC . ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
CC MAFK.
CC
CC -1- SUBUNIT: HETERODIMER OF BACH2 AND MAFK.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
CC NEURONAL CELLS.
CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC -----
CC

[illegible]

Db 237 CTKNVYSAPSHGTSFASTFSDSPGNSLKPLGPMGIKSEPPSEETEEITLCLSGDE 296
 Qy 125 TIWQIKGKPOVIYQETTDPI-----FLMNEKGMLTK 157
 Db 297 T---DIKORPGDVENDRKQSPARTPTRTGACILDRSRSVSSPCLSLFSLFITKGV--- 350
 Qy 158 KDAKOLEYAAKQFTPLSLF-----DLDRNTPLWSS--RPHNPMYVLPFIMHG 204
 Db 351 -ESTGLPSTSQ--PLVRSSACPFNKISQGLKTDYPLAGNYGPHVGQKDVSNFAMG 407
 Qy 205 KPNRSP-----NTSHEAKQFTPEFRAPELK-----FQSVKVKAAEDLNGTSD 250
 Db 408 SPLRPGPETICEFSSSPCSQAGRELATEHQEPGLMGDMQVNRPOIK-CESYGTNS 466
 Qy 251 LWFQYTOQSHWQIFNGKNSRPRVRHYDQPEIFLTQPV--YSDLPWDGKVRMIGMCAVHHS 308
 Db 467 DESG-----SFSEADSESCVPDQGEVKLPFPVQITDLPNDQFMKMK---HKL 515
 Qy 309 NGESAKL-----SRSMNR 321
 Db 516 TSEQLEFIHDIRRSKNR 533

RESULT 11
 SRK6_BRAOL STANDARD; PRT; 849 AA.
 AC Q09092;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Brassica.
 ON NCBI_TaxID=3712;
 RN SEQUENCE FROM N.A.
 RP STRAIN=CV. S6S6; TISSUE=Stigma;
 RC MEDLINE=92020942; PubMed=1681543;
 RA Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RT "Molecular cloning of a putative receptor protein kinase gene encoded
 RT at the self-incompatibility locus of *Brassica oleracea*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -!- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 CC (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CC CYTOPLASMIC DOMAIN.
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PILIST AND ANTHER.
 CC -!- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 CC B.OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -!- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M76647; AAA33000.1; ALT_TERM.
 CC HSP; P11362; IFGI.
 CC InterPro; IPR001480; B_lectin.
 CC InterPro; IPR000719; Euk_kinase.

DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR002290; Ser_thr_kin_actsite.
 DR InterPro; IPR000858; Slocus_glycop.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00954; S_locus_glycop; 1.
 DR SMART; SM00108; B_lectin; 1.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; Signal; ATP-binding;
 KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32
 FT CHAIN 33 849
 FT PUTATIVE SERINE/THREONINE KINASE
 FT RECEPTOR.
 FT DOMAIN 33 446
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 447 466
 FT DOMAIN 467 849
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 528 779
 FT PROTEIN KINASE.
 FT NP_BIND 534 542
 FT ATP (BY SIMILARITY).
 FT BINDING 556 556
 FT ATP (BY SIMILARITY).
 FT ACT_SITE 653 653
 FT CARBOHYD 47 47
 FT CARBOHYD 120 120
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 196 196
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 260 260
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 314 314
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 389 389
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 442 442
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 849;
 Best Local Similarity 21.6%; Pred. No. 13;
 Matches 77; Conservative 51; Mismatches 107; Indels 121; Gaps 24;

Qy 105 LVHGETPAVTKRSIRLDETIWQIKGPOVIYQETTDPIFLMG-----NEK---GMLTK 157
 Db 24 LIHPALSIYINT-----LSSTESLTISSNKTIV---SPGSIFEVGFRTNSRWYLGMYK 75
 Qy 158 K-DAKOLEYAAKQFTPLSLFSD-----LDNRNTPWSSRPHNPMYVLPFIMHGK 205
 Db 76 KVSDRTYVWVANRDNPLSNAIGTLKISGNNLVLLDHSNKPVMWTN-----LTRGN 125
 Qy 206 PNRSP-----NTPSHEAKQFT-----PNEFRAPELKQVSVKVKAAEDLW 245
 Db 126 -ERSPVVAELLANGFNVRDSSNNDASEYLWQSFYDPTDTPLLPENKLGYNLKT----- 177
 Qy 246 GTDSDLWFGYTOQSHWQIFNGKNSRPRVRHYDQ-----PEIFLTQ---PVYSDLPWDGK 296
 Db 178 GLNREL-----TSWRSSDDPSSGNF---SYKLETQSLPEFYLSRENFPMHRSGPWNG- 226
 Qy 297 VMTGMGAVH-----HSNGESAKLSRWNRAY-----LMAGMEWKNLTVMV-RI 339
 Db 227 IRFSGIPEDQKLSVMYNFNIENNEEVAYFTFTMNNFSRLTLISEGYFQRLTWPSIRI 286
 Qy 340 WGRIFKEGSGOPDNDPILDY-----YGYDVRFLYOLENKNKNSISGTVR-YNPRS 389
 Db 287 WNRFW-----SSPVD-POCDYIIMCYPAYCDV-----NTSPVCNCIOGFNPRN 329

RESULT 12
 YL70_ARCFU STANDARD; PRT; 324 AA.
 AC O28112;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHEITICAL PROTEIN AF2170.
 GN AF2170.
 OS Archaeoglobus fulgidus.
 CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;

```

OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Krelavage A.R., Graham D.E., Kyrplides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kane B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: CONTAINS 2 KELCH REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000955; AAB89093.1; -.
DR TIGR; AF2170; -.
DR InterPro; IPR001798; Kelch.
DR Pfam; PF01344; Kelch; 2.
KW Hypothetical protein; Repeat; Complete proteome.
FT REPEAT 229 276
FT REPEAT 277 323
FT REPEAT 324 360; C445388CFEB96E45 CRC64;
SQ SEQUENCE 324 AA; 36025 MW; C445388CFEB96E45 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 324;
Best Local Similarity 20.1%; Pred. No. 3.7;
Matches 54; Conservative 44; Mismatches 72; Indels 98; Gaps 18;

QY 224 EFRAPELKQVSVKVKAAKD-----LWGTDSDFQREANVWDGGEILFGT-----VFENGKYSPTDQIL 112
DB 63 EHFPPN-RLEI---LLSSEDFQREANVWDGGEILFGT-----VFENGKYSPTDQIL 112
QY 276 DYQPEIFLTQPYVSDLP-----W-DGKVRMI-----GMGAVHSHNGESAKLSRS 318
DB 113 SPNPKLERLRVLNASLPHPTSDVAAVWGDSDRVYIFLNNSECEVYAFYPSNESFAKLDSV 172
QY 319 -----W-NRAYL-----MAGMEKNTVMPRIWGRIFKEGSGSQ 351
DB 173 CPTEHPGCVSHVWYGGKAYFCGEGVASFDPMGFKW--IAFTDRVWVRAATVADG-- 228
QY 352 PDNDPDLIDYGYGDRVFLYQLENKSNISGT---VRYNPRSGKGAOLDYVYPLGKG-- 405
DB 229 -----YFAFGSSGTAETKDEIRFPNPKTGE-LCEMRTKLPVARGQA 270
QY 406 --ISG-YFOIF-----QYQGSLIDYNH 425
DB 271 VAVGGEIYIFGGYTKDGYVANEILRDYD 298

RESULT 13
R114_HUMAN
ID R114_HUMAN STANDARD; PRT: 1158 AA.
AC P48552;
DT 01-FEB-1996 (Rel. 33, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NUCLEAR FACTOR RIP140 (NUCLEAR RECEPTOR INTERACTING PROTEIN 1).

```

```

GN NR1P1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RX TISSUE=Breast;
RX MEDLINE=95369246; PubMed=7641693;
RA Cavailles V., Dauvois S., L'Horsset F., Lopez G., Hoare S.,
RA Kushner P.J., Parker M.G.;
RT "Nuclear factor RIP140 modulates transcriptional activation by the
RT estrogen receptor."
RL EMBO J. 14:3741-3751(1995).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Sharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
CC -!- FUNCTION: MODULATES TRANSCRIPTIONAL ACTIVATION BY THE ESTROGEN
CC RECEPTOR.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X84373; CAA59108.1; -.
DR EMBL; AL163207; CAB90396.1; -.
DR MIM; 602490; -.
KW Transcription regulation; Nuclear protein.
FT CONFLICT 124 124 P -> R (IN REF. 1).
FT CONFLICT 448 448 R -> G (IN REF. 1).
FT CONFLICT 721 726 NRGKSE -> TGRVK (IN REF. 1).
FT CONFLICT 1080 1080 T -> A (IN REF. 1).
SQ SEQUENCE 1158 AA; 126941 MW; 81FC424968E9A5F6 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 1158;
Best Local Similarity 19.3%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 150; Indels 214; Gaps 26;

QY 3 VSLSTLTLILSCFALLAQQAQVNPVAFVDEVESEND-----LQDNE 48
DB 458 VSLDNFTQSLNTW-----DPKVPDVIKEDQDTSKNLSHOKVTLQLLLGHKNE 510
QY 49 LPIDVQSATQ-----SASTDTANPLDEHEPELYTTA-----L 80
DB 511 ENVEKNTSPQGVHNDVSKENTQNYARTSVIESPSTRTTPVS--TPPLTSSKAGSPINL 568
QY 81 ENKTMILN-----CSALNQDIMRLACYDTL-----VHGETPAVVK-----TKRSI 120
DB 569 SQHSLVKWNSPPYVCSTQSEKLTNTASNHSMDLTLSKDPKPPKPAQNEGAQNSATFSAS 628
QY 121 RLDETWTQIKGPQVIYQETTDPIFLMGNEKGMCLKAKQLEYAAKQFTPLSLSLFDLD 180
DB 629 KLLQNLQAQCGMQSSMSVSEQRPFSKQLLTGN-----TDK-----PIGM---ID 667

```

Qy	314	KLRSRNRAYL	-----MAGMENKNTVMPRIWGRIFKGGSGQPDNP-----	350
		: : : : :	: : : : : : : : :	
Db	821	RLLRQNQDSYLADDS	DRSHRNEMALLESKNLCWPK-----KRKLYTEPLENFKKMK	874
Qy	357	-DILDYVG-----	YGDVRFYQLENKSNISGTVRYNPRSGKGALQDYVYPLGK	405
		: : : : :	: : : : : : : : :	
Db	875	NNTVDAANHNSAPEVLY	SGSLNNOELKFSRN-----DLEFYKYPAGHG	916

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

Query Match 3.9%; Score 91.5; DB 1; Length 467;
Best Local Similarity 19.8%; Pred. NO. 6.8;
Matches 62; Conservative 44; Mismatches 128; Indels 79; Gaps 14;
OV 134 POVIY-----OETTDPIFLMG-----NEKGMITKKDAKOLEVAAKOFTEPLSLSFOLDLN 182

[illegible]

Qy 421 -----IDYNH 425
:|:|
Db 417 EKASREQOLDYSH 429

Search completed: November 30, 2001, 14:27:00
Job time: 547 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:45 ; Search time 36.79 Seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-6
Perfect score: 2360
Sequence: 1 MKVSLSTLTLILSCFAILA.....YNHEATSGVGLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.5	4.3	682	3	US-08-481-435-6
2	100.5	4.3	1726	2	US-08-609-049A-30
3	100.5	4.3	1726	4	US-09-170-996-30
4	98.5	4.2	1658	2	US-08-609-049A-13
5	98.5	4.2	1658	4	US-09-170-996-13
6	90.5	3.8	1024	4	US-09-091-117-5
7	89.5	3.8	857	1	US-07-717-331F-2
8	88	3.7	537	2	US-08-633-879C-2
9	87.5	3.7	503	1	US-07-946-497-2
10	87.5	3.7	503	1	US-08-483-322-2
11	87.5	3.7	503	2	US-08-478-882-2
12	87.5	3.7	666	4	US-08-961-083-2
13	86	3.6	781	1	US-08-373-134D-2
14	86	3.6	781	2	US-09-114-637-2
15	86	3.6	1088	3	US-08-633-768A-1
16	85	3.6	816	1	US-07-731-157A-4
17	85	3.6	816	1	US-08-229-444B-2
18	85	3.6	816	2	US-08-541-780-4
19	85	3.6	985	5	PCT-US96-03916-6
20	85	3.6	985	5	PCT-US96-03916-66
21	84.5	3.6	774	3	US-08-902-632-2
22	84.5	3.6	774	3	US-09-073-354-1
23	84.5	3.6	774	3	US-08-656-005A-1
24	84.5	3.6	774	4	US-09-073-209-1
25	84.5	3.6	774	4	US-09-363-095-1
26	84.5	3.6	774	4	US-09-418-027-1
27	84	3.6	522	6	RE34606-6

Query Match 4.3%; Score 100.5; DB 3; Length 682;
Best Local Similarity 20.4%; Pred. No. 0.11; Indels 103; Gaps 24;
Matches 92; Conservative 68; Mismatches 189;

ALIGNMENTS

RESULT 1
US-08-481-435-6
; Sequence 6, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 682 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-481-435-6

Query Match 4.3%; Score 100.5; DB 3; Length 682;
Best Local Similarity 20.4%; Pred. No. 0.11; Indels 103; Gaps 24;
Matches 92; Conservative 68; Mismatches 189;

Sequence 9, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 145, App
Sequence 48, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 4, Appli

```

US-08-609-049A-30
Query Match 4.3%; Score 100.5; DB 2; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.54;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

QY 1 MKVSL-----TLTSLSCFAILAIQAQAKVNPVAFVD-----EVRSENDIGQD 46
   :|:|:|
Db 465 VKVSEIEGLQPLPFTTCDVSSVTEIIMQALSWHDLNQVDVGSYILKVGQEEVLQN 524
   :|:|:|
QY 47 NELPIDVQSQTASTDTANPLDHEPELVTALENKTMILNCSALMODIMRLACYDTLV 106
   :|:|:|
Db 525 NHCLGSHEHQNCRKWD-----EIKLQLLTLSAMCONLARTAEDD--- 565
   :|:|:|
QY 107 HGETPAVVKTKRIRLDETITWQTIKGPQVYQETTDPIF----- 146
   :|:|:|
Db 566 --EAP-----VDLNKYLQYIEKPYKEVMTHPVBEELDSYHYQVELALQTNQHRV 615
   :|:|:|
QY 147 -----LMGNEKGMLTKKDAQLEYAA-----KQFTPLSLSFDLDRNNTPLWS 188
   :|:|:|
Db 616 DOVTKAVRKTCSDLDGVETSVTEA-VKKLRAVLNPRKNSADVTSLSGSDTRKNSTKGS 674
   :|:|:|
QY 189 SRPNPMYVLP-----IFMHGPNR-----SPNTPSHEARQFTPNFRAPELKFQV 234
   :|:|:|
Db 675 LNPNPQVSMDDLTHTRIYDLRLRHANSRSCSTGCPGRSRIKEAWTATE-----QL 726
   :|:|:|
QY 235 SVKVKAAEDL---WGTUSDLDWFGYTQSHHQWIFNGKN-SRPF---RVHDYQPEIFLTQ-- 285
   :|:|:|
Db 727 QFTVYAAHGIGSSNVSNYKYYLICLSH-----NGKDLFKPIQSKKYGTGKNFYLIKWD 782
   :|:|:|
QY 286 ----PV-YSDLPWDGKVRMIGMAVHNSNGESAKLRSWN-----RAYL 324
   :|:|:|
Db 783 ELIIFPTQISQPLSVLHLTFLGVLNQSSGSDPSNKKQKGPALGKVSILTLDFKRFL 842
   :|:|:|
QY 325 MAGMEWKNLWMPRIWGRIPKESGSDPDNDPILDIYGYGDVRFVLENKSNISGTVR 384
   :|:|:|
Db 843 TCG-----TKLLYLW-----TSSHTNSIPGAIPKSYVMERIVLQVDVFPSPAFDIY 889
   :|:|:|
QY 385 YNPESGKALQDQVYPLGKIGISGYFOIFQYGGSLDYNHEATSFVCV 432
   :|:|:|
Db 890 TSPQIDRNIQQDKLETLESIDK-----KLLDIHRSDFGL 927
   :|:|:|

RESULT 3
US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996

```


APPLICANT: Williams, Lewis T.
APPLICANT: Molz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0637000S
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.2%; Score 98.5; DB 4; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.81;
Matches 101; Conservative 78; Mismatches 195; Indels 159; Gaps 23;

Qy 5 LSTLTLSLSCFAILAIQ---QAKAVPNPVPFVDEVR-----ENDLGQDNE 48
Db LSPVTQVRNMGENASVKVSIIEGLQLPVFTCDVSTVEIIMQALCWVHDDLQ--- 437
Qy 49 LPIDVQS-----ATQASSTDITANPLDEHE--PELYTTALENKTMLINGSALNQDIMRLAC 101
Db --VDVGSYILKVGQEEVLQNNHCLSGSHEHQNCRKWDEIKLQLLTLSAMCQNLARTAE 495
Qy 102 YDTLVHGETPAVTKRSIRDETQWIKGPKQPVVYQETDPIF----- 146
Db -----EAP-----VDLNKLYIQEKPYKEMIRHPVEELDSYHYQVELALQFEN 542
Qy 147 -----LMGNEKGMKTKDAKOLEYAA-----KQFTPLSLSPFLDRNN 183
Db 543 QHRAVDQVIKAVRKCSALDGVETSPVTEA-VKLLKRAVNLPRNKSADVTSLSGSDTRKN 601
Qy 184 TPLWSSRRPHNPMYVLP-----IPMHGKPNR-----SPNTPSHEARQFTPNNEFRAPE 229
Db 602 STKGSLENPNVQVSMDDLTTAIVDLLRLHANSSRCSTGCPGRSRIKEAWTATE----- 656
Qy 230 LKFOVSVAKKAEDL---WGTDSDLWFGYTOQSHWQIFNGKN-SRPF---RVHDIQPIEF 282
Db 657 ---QLQFTVYAAHGSSWNWVNYEKYILICLSH-----NGKDLFKPIQSKKVGTYKNFFY 709
Qy 283 LTQ-----PV-YSDLPDGDKVRMIGMGAVHHSNGESAKLSRSWN----- 320
Db 710 LIKWDELIIFIQISQLESVILHTLFGVLNQSSGSPDSKQKRGPEALGKVSILTIFD 769

Qy 321 -RAYLMAGMEWKNLTVMPRINGRIFKEGSGSQPDNDPILDYGYGVDYRFLYQLENKSN1 379
Db 770 PKRFLTCG-----TKLLYLW-----TSSHTNSIPGAIPKSVYMERIVLQVDFPSPA 816
Qy 380 SGTVYNPRSGKGLQLDYVYPLKKGISGYQIFQGYGQSLIDYNHEATSGCV 432
Db 817 FDIYTSPOIDRNIIQDKLETLESIDKG-----KLLDIHRDSSFG 859

RESULT 6
US-09-091-117-5
Sequence 5, Application US/09091117
Patent No. 6171589
GENERAL INFORMATION:
APPLICANT: The University of Melbourne
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
TITLE OF INVENTION: Vaccines
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,117
FILING DATE: 12 JUNE 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00803
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7127
FILING DATE: 13-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: WINNER, Ellen P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: +1 303 499 8080
TELEFAX: +1 303 499 8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1024 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mycoplasma genitalium
US-09-091-117-5

Query Match 3.8%; Score 90.5; DB 4; Length 1024;
Best Local Similarity 20.2%; Pred. No. 2.6;
Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

Qy 1 MKVS-LSTLTLSLSCFAILAIQAKAVPNPVPFVDEVRSENDLGQDNELP1-DVQSATQ 58
Db 1 MKLSTITTTICLSISGAFCTTAI---ALPTTVALLKNHQOQNTKQKN--PIKDIRFGLN 54
Qy 59 SASDITANPLDEHEPELYTTALENKTMLINGSALNQDIMRLACYDTLVHGETPAVTKR 118
Db 55 NVQVPNTIPL--HQTVVEVT--NNKAIYDYKDAKQKFFL-----AKSALNNKL 98
Qy 119 SIRLDETQWIKGPKQPVVYQETDPIFLMGNEKGMKTKDAKOLEYAAKO-FTPLSLSF 177
Db 99 QVEFDKFLLR-----GVINALNADLKKEWIDQTLFIPNQSF 135
Qy 178 DLDRNNTPLWSSRRPHNPMYVLP-IFMHGKPNRSP-NTPSHEARQFTPNNEFRAPELKFQVS 235

Db 136 DLSANKNLTLNSQSEVSLDLLEIFTNFSDKNQPLKLPEDGCVVNVANE-----SYTYSVK 191
QY 236 VKYAAEDLWGTDSDLWFGYTOQSHWOIFNGKNSRPFVRVHDYOPEI-FLTQPVYSDL-PW 293
Db 192 ATLQKLKLVTSRADHSVGSYAIPVSLNGKTQDNFSENPFSNINFAFKVYVNALNPF 251
QY 294 DGKVRMIGMAVHVS--NGESAK-----LSRSWNRAYLMAGMEWKNLTVMPWIRGRIK 345
Db 252 EAQOYLVGQCKFLNQKVNADDVKNDINNHIETQFNAKITA-----TLGKAFK 300
QY 346 ---EGSGSOP-----DNDPILDVY--GYGDV-----REFYQL-- 373
Db 301 QFGEHNGOPLSLKLVSLGNLNEFKQLFNVRPGLGDFVSLDIQSSQSSNKKTVIQLLF 360
QY 374 ENKSN1 379
Db 361 ENKTTI 366

RESULT 7
US-07-717-331F-2
; Sequence 2, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incomparability Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-717-331F-2

Query Match 3.8%; Score 89.5; DB 1; Length 857;
Best Local Similarity 21.4%; Pred. No. 2.4;
Matches 76; Conservative 50; Mismatches 110; Indels 119; Gaps 23;

QY 105 LVHGETPAVTKRSRLDETWTQTKGPQVYQYTTDPILFMG-----NEK---GMLTK 157
Db 24 LTHPALSIYINT-----LSTESTLTISSNKTLY---SPGSIIEVGFRTNSRWYLGMYK 75
QY 158 K-DAKOLEYAAKQFTPLSLFSD-----LDKNNTPLASSRPHNPVYLPPIEMHCK 205
Db 76 KVSDDRYVWVANDNPLNSAIGTLKISGNLVLLDHSNKPVMWTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEARQFT-----PNEFRAPELKFQSVKVKAAEDLW 245

Db 126 -ERSPVVAELANGNFVMRDSSNNDASEYLWQSFDPYPTDTLLPEMKLYGLKT----- 177
QY 246 GTSDLDWFGYTOQSHWOIFNGKNSRPFVRVHDYO-----PEIFLTQ---PVYSDLPWDGK 296
Db 178 GLNRFL-----TSNRSSDDPSGNF---SYKLETSQSLPEFYLSRENFMHRSRGPWNG- 226
QY 297 VRMIGMAVH-----HNGESAKLSRSWNRAY-----LMAGMEWKNLTVMP--RI 339
Db 227 IRESGIPEDQKLSYMYVNFETENNEEVAYTFRTMTNNSFYSLTLISEGYFORLTWYPSIRI 286
QY 340 WGRIFREGSGSQPDD-----NPDILDYGYGDVRFYQLENKNSISGTVR--YNPRS 389
Db 287 WNRFSWSPVDROCDTYMCGP-----YAYCDV-----NTSPVCNIOGFNPRN 329

RESULT 8
US-08-633-879C-2
; Sequence 2, Application US/08633879C
; Patent No. 5928922
; GENERAL INFORMATION:
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Pihlajaniemi, Taina
; APPLICANT: Helaakoski, Tarja I.
; APPLICANT: Annunen, Pia P.
; APPLICANT: Nissi, Ritva K.
; APPLICANT: No. 5928922elainen, Minna K.
; TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
; TITLE OF INVENTION: METHODS FOR PRODUCING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,879C
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0041-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 537 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-633-879C-2

Query Match 3.7%; Score 88; DB 2; Length 537;
Best Local Similarity 21.4%; Pred. No. 1.6;
Matches 100; Conservative 52; Mismatches 145; Indels 170; Gaps 26;

```

QY 1 MKVSLSTLTSLSCFAILAIQAKAVPNPFAVDEVRSENDLQGDNELPIDVQSA---- 56
DB 1 MKQLVLVLV-LMSWFGVLSWQAEFFTSIGHTMDLIYAACKDLVOSLKEYILVEAKLAK 59
QY 57 -----TQASTD-----TANPLDEHE-----PELYTALENKT--MLINCS 90
DB 60 IKSWSKMEALTSRSAADPEGVLAHPVNAVYKLVKRLNTDWPALGDLVLQDASAGFVANLS 119
QY 91 ALNQ-----DIMRLA-CY-----DVLVHGETPAVIKTKRSIRLDETI----- 126
DB 120 VQRQFPFDEDESGAARALMRLODYKLDPDITISGELPGT-KYQAMLSVDCFCGLGRSA 178
QY 127 -----WOTIKGPOVYQETTDPIFLMGNEKMLTKK--DAKQLEYAAKQFTPLSLSFDL 179
DB 179 YNEGDIYHTLVMEQVLQ-----LDAGEBATVTKSLVDLYSAVAVQLGDLHRAVEL 231
QY 180 DRNNTPLSSRPHNPMYVLPFPMHGKPNRSPNTPSHE-----ARQFTP--NEFRAPELKF 232
DB 232 TRRLLSL-----DPSHERAGGNLRYFERLLEERKSLSN 266
QY 233 QVSVKVKAADLWGTDSLWFGYTOQSHWQIFGNKNSRPFVHDYQPEIFLTQPVYSDLP 292
DB 267 QTDAGLATQENLY-----ERP-----TDLPE-----RDVYESL- 295
QY 293 WDGKVRMIGMGAVHSHNGESAKLS-RSMNRAYLM--AGMEWKNLTVMPRIMGRIFKEGSG 349
DB 296 -----CRGEGVLTERRQKKLPCRYHGHGVRVPLLIAPIAP-----EKE----- 331
QY 350 SQPDNDPILDYGGDVRFYQLENKSNIS-----GTVRYNPRSG 390
DB 332 EDEWDSPHIVRY-----DVMSDEIEREIKAIKPKLARATVR-DPKTG 374

```

RESULT 9

```

US-07-946-497-2
; Sequence 2, Application US/07946497
; Patent No. 5506119
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946.497
; FILING DATE: 19921109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids

```

```

; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-946-497-2

Query Match 3.7%; Score 87.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. NO.1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDLQGDNELPIDVQSAQTSASTDTANPLDEHEPELYTALENKTMLINCSALNQDIMRL 99
DB 169 QEDIDASNIIDVSGS-----GSTIEKSTPEGYILHTDLP-----SQPTGDRD----- 212
QY 100 ACYDTLVHGETPAVIKTKRSI-----RLDETITWOTIKGPOVYQETTDPIFLMGNEKG 153
DB 213 ---DAFFGTSLTATATTPWVSAHTKQKQRTQWNPPIHNSPEVLIQTTT----- 258
QY 154 MLTKKDAKQLEYAAKQFTPLSLSFOLDNRNTP-----WSSRPHNPMYVLPFPMHGKPNRS 209
DB 259 -----RMT-----DIDRNSTSAHGENWTQEPQPP----- 283
QY 210 PNTPSHEARQFTPNFRAPELKFQVSVKAAEDLW-----GTDSDLWFGYTOQS 259
DB 284 -NNHEYQDEEETPH-----ATSTTADPNSTTEEAATQKEKF-----EN 322
QY 260 HWOIFENGKN-----SRPFVHDYQPEIFLTQPVYSDLPW-----DGKVRMIGMG 303
DB 323 EQW---GKNPTPSEDSSHVTEGTASAHNNHPSQRMVTSQEDVSWTDFDFDISHPMGOG 379
QY 304 AVHSHNGESA 313
DB 380 HOTESKGHSS 389

RESULT 10
US-08-483-322-2
; Sequence 2, Application US/08483322
; Patent No. 5760178
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483.322
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946.497
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5300

```

TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-322-2

Query Match 3.7%; Score 87.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDLQDNELPIDVQSATQSASTDTANPLDEHEPELYTTALENKTMILNCSALNQDIMRL 99
Db 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYILHTDLP---SQTGDRD--- 212
QY 100 ACYDTLVHGETPAVTKRSI-----RLDETIVQIKGPKQVYQETTDPIFLMGNEKG 153
Db 213 ---DAFFIGSTLATIATTPWVSAAHTKQNCQERTQWNPISNPEVLLQTTT----- 258
QY 154 MLTKKDAKQLEYAAKQFTPLSLSFDRNTPL-----WSSRPHNPMVLPPIFMHGXPNRS 209
Db 259 -----RMT-----DIDRNSTSAHGENWTOEQPPF----- 283
QY 210 PNTPSHEARQFTNEFRAPELKFQSVKVAEEDLW-----GTDSDLWFGVYTOQS 259
Db 284 -NNHEQDEEETH-----ATSTWADPNSTTEEAATQEKWF-----EN 322
QY 260 HMQIFNGK-----SRPRVHDYQPEIFLTQPVYSDLPW-----DGKVRMIGMG 303
Db 323 EMQ---GKNPPTPSDSHVTEGTTASAHNNHPSQRTTOSQEDVSWTDFDPIISHPMGQG 379
QY 304 AVHHSNGESA 313
Db 380 HOTESKGHSS 389

RESULT 11
US-08-478-882-2
Sequence 2, Application US/08478882
Patent No. 5885575

GENERAL INFORMATION:
APPLICANT: HERRLICH, Peter
APPLICANT: PONTA, Helmut
APPLICANT: GUENTHER, Ursula
APPLICANT: MATZKO, Siegfried
APPLICANT: WENZL, Achim
TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
TITLE OF INVENTION: SEQUENCES CODING THESE ANTIBODIES AGAINST THESE PROTEINS,
TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478.882
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/946.497
FILING DATE: 19921109
ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16915/145
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-882-2

Query Match 3.7%; Score 87.5; DB 2; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDLQDNELPIDVQSATQSASTDTANPLDEHEPELYTTALENKTMILNCSALNQDIMRL 99
Db 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYILHTDLP---SQTGDRD--- 212
QY 100 ACYDTLVHGETPAVTKRSI-----RLDETIVQIKGPKQVYQETTDPIFLMGNEKG 153
Db 213 ---DAFFIGSTLATIATTPWVSAAHTKQNCQERTQWNPISNPEVLLQTTT----- 258
QY 154 MLTKKDAKQLEYAAKQFTPLSLSFDRNTPL-----WSSRPHNPMVLPPIFMHGXPNRS 209
Db 259 -----RMT-----DIDRNSTSAHGENWTOEQPPF----- 283
QY 210 PNTPSHEARQFTNEFRAPELKFQSVKVAEEDLW-----GTDSDLWFGVYTOQS 259
Db 284 -NNHEQDEEETH-----ATSTWADPNSTTEEAATQEKWF-----EN 322
QY 260 HMQIFNGK-----SRPRVHDYQPEIFLTQPVYSDLPW-----DGKVRMIGMG 303
Db 323 EMQ---GKNPPTPSDSHVTEGTTASAHNNHPSQRTTOSQEDVSWTDFDPIISHPMGQG 379
QY 304 AVHHSNGESA 313
Db 380 HOTESKGHSS 389

RESULT 12
US-08-961-083-2
Sequence 2, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-2

Query Match 3.7%; Score 87.5; DB 4; Length 666;
Best Local Similarity 20.1%; Pred. No. 2.6; Mismatches 188; Indels 97; Gaps 24;
Matches 88; Conservative 64; Mismatches 188; Indels 97; Gaps 24;
QY 25 KAVPNPFAFVDEVRSENDL-GODNELPIDVQSATQS-----ASTDTANPLDEHEPE 74
DB 1 KIYDNKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRLGAFLRN 60
QY 75 LYTTALENKTMLNCSALNQDIMRLACDYLTVHGETPAVTKRSIRLDETITWQTIKGP 134
DB 61 LOSNSLOGG-----STLTQOLIKLTYESTSTSDQIS-----RKAQEAWLAIQLEQ 106
QY 135 OVVOYQET-----TDPIFLMGNEKGLTK-----DKAKOLE-----YAAKQFTPL 173
DB 107 KATQKEILTYINKYMGNGVGMOTAAQNYGKDLNLSLPQALLAGMPQAPNQIDPY 166
QY 174 S-LSPDLRNNTPMLSSRPHNPVLPILFMHGKPNRSPNTPSHEARQFTPNEFRAPE--- 229
DB 167 SHPEAAQDRRLVL--SEMKNOGYI---SARQYKAVNTPITDGLQSLKSASNPAYMD 220
QY 230 --LKFQVS-VKVAEADLWGTDSLWFGYTO--QSH-WQIFNGKNSRPFVHDYQPEIFL 283
DB 221 NYLKEVINQVEETGYNLLTTGMVYTNVDQAQKHLWDIYNTDEYVAYPDDELQ----- 275
QY 284 TQPVYSDLPWCKVRMIGMGAHVHNS-----GESAKLSRSNNRAYLMAGMEWKNLT-VM 336
DB 276 VASTIVDVS-NGKV-IAQLGARHQSNNVSFGINQAVENRQW-----GSTMKEFITDYA 326
QY 337 PRI-----WGRIFKEGSGQPDNDILDY-YGY-GDVRFLYQLENKSNISGTVRYN- 386
DB 327 PALEYGVYDSTATIVHDEPYNPYPTNPVYNWDRGYFGNITLQYALQOSRNVPAVETLNK 386
QY 387 -----PRSGKGAQLQDY 398
DB 387 VGLNRAKTFLLGLGIDY 403

RESULT 13
US-08-373-134D-2
Sequence 2, Application US/08373134D
Patent No. 5780296
GENERAL INFORMATION:
APPLICANT: Kmiec, Eric
APPLICANT: Holloman, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/114,637
FILING DATE:

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/373,134D
FILING DATE: January 17, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Friebe, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 7991-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 781 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-373-134D-2

Query Match 3.6%; Score 86; DB 1; Length 781;
Best Local Similarity 20.7%; Pred. No. 4.9; Mismatches 122; Indels 32; Gaps 9;
Matches 50; Conservative 37; Mismatches 122; Indels 32; Gaps 9;
QY 39 SENDLGOONELPIDVQS-ATQASDSTANPLDEHEPELYTTALENKTMLNCSALNQDIM 97
DB 533 ARNDGTSQAPTSAFSGTGSALPDQPLAMDVASQTAFSTCLASIAPTAEAVGAREL 592
QY 98 RLACYDTLVHGETPAVTKRSIRLDETITWQTI-----KGPQVYQY-----TTDPI 145
DB 593 DSAC-----ASNDVPLRTLEARTAQLGQT-WSNLIINVFLSKTRARICMRDDQAPACEPV 647
QY 146 FLMGNEKMLTKKDAKOLEYAAKQFTPLSLF-DLDRNNTPLSSR-----PHNPMVLP 200
DB 648 RONTNQRTASKSLMNTYRKAAVINPFQATMLDVGVDKSAQLRQFVITPRKAVHVL-- 705
QY 201 FMHGKPNRSPNTPSHEARQFTPNEFRAPELKFQSVKVAEADLWGTDSLWFGYTOQSH 260
DB 706 -----NAYPSTVMH-AMHATADSTPAPESSQOQRAAERHPAEQEDADQDLFGEALQEH 758
QY 261 W 261
DB 759 W 759

RESULT 14
US-09-114-637-2
Sequence 2, Application US/09114637
Patent No. 5945339
GENERAL INFORMATION:
APPLICANT: Kmiec, Eric
APPLICANT: Holloman, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/114,637
FILING DATE:


```
; APPLICATION NUMBER: US/08/633,768A
; FILING DATE: 02-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9321301.5
; FILING DATE: 15-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: DYOUT.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-633-768A-1

Query Match      3.6%; Score 86; DB 3; Length 1088;
Best Local Similarity 21.6%; Pred. No. 8.4;
Matches 72; Conservative 40; Mismatches 138; Indels 84; Gaps 15;

Qy   106 VGHETPAV-IYKRSIRLDETIWITIGKPQQVVQEITDPIFLMGNEKGMLTKDKDAQLE 164
     :| |||| |:::||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    29 IHSAFPVAHTATRTKNRLN--VSMTALSDKQTATAGSTD-----NPDMI----DYKYTD 76

Qy   165 YAAK-QFTPLS-----LSFDLDNRNTPLWSSRPHNMVYLPIFM 202
     |:|||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    77 YGVWGFSPLTNWFNAAGSTPGGITDWATMNVNFDRIDNP--SIVQHVPVQOVTSY 134

Qy   203 HGKPNRPENTPSHEARQFTPNEFRAPELKFQSVVKVKAADLMGTDSDLMFEGYTQQSHWQ 262
     :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db   135 NNSVYRVRFDPGPIRDVT---RGPILKQLD-WIRTQELSEGCDCPM-----178

Qy   263 IFNGKNSRPFRVDHYOPEIFLTQPYSDDLPWDGKVRMIGMCVAVHSGESAKLSRWRA 322
     |:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db   179 TFTSEGLTFEETKLSQLSVIIYGFKTRVRTRKSDGKVM-----ENDEVGTAS-----224

Qy   323 YLMAGMEWKNLTVMPRIWGRIFKEGGSGSQPDNDPDI LDYYGYGDVRFLYQ---LENKSNI 379
     :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db   225 ---SONKCRCGLMWFDRLIGNAIASNKNFRDNADVKBGEFGAGEVNCKYQDTYILERTGI 281

Qy   380 SGT-----VRYN-----PRSGKGALQLDVYVL 402
     :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db   282 AWTNYNDLNYNQNWLRRPPHHDGALNPDDYIPM 315

Search completed: November 30, 2001, 14:17:48
Job time: 225 sec
```

Search completed: November 30, 2001, 14:17:48
Job time: 225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:16:59 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHEATSGVGLMLNDWMGL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*

1:	/SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	2360	100.0	442	21	AAV85270
2	2352	99.7	442	21	AAV85268
3	2348	99.5	442	21	AAV85271
4	2330	98.7	442	21	AAV85269
5	822	34.8	370	21	AAV75156
6	815	34.5	370	21	AAV75157
7	815	34.5	370	21	AAV75158
8	815	34.5	370	21	AAV70629
9	812	34.4	375	21	AAV70628
10	246.5	10.4	355	19	AAW98871
11	239.5	10.1	356	19	AAV10960

12	156.5	6.6	253	18	AAW20760
13	114	4.8	1686	19	AAW70991
14	109.5	4.6	800	8	AAV70420
15	109.5	4.6	822	13	AAV26021
16	108	4.6	824	21	AAV23180
17	107	4.5	157	18	AAW20538
18	104.5	4.4	537	22	AAV91000
19	104	4.4	798	21	AAV40925
20	104	4.4	798	21	AAV90225
21	104	4.4	798	22	AAV83944
22	100.5	4.3	682	17	AAW04359
23	99.5	4.2	1726	18	AAV38756
24	98.5	4.2	467	20	AAV13378
25	98.5	4.2	467	21	AAV94860
26	98.5	4.2	467	22	AAV88401
27	98.5	4.2	467	22	AAV80246
28	95.5	4.0	502	22	AAV90031
29	95.5	4.0	516	22	AAV78876
30	95	4.0	522	22	AAV95616
31	94	4.0	372	16	AAV69607
32	94	4.0	761	20	AAV99084
33	94	4.0	1115	21	AAV93947
34	93.5	4.0	467	19	AAV75057
35	92.5	3.9	1717	22	AAV20498
36	92	3.9	888	22	AAV70751
37	91.5	3.9	857	13	AAV29814
38	91	3.9	1227	22	AAV81501
39	90.5	3.8	1024	18	AAV19604
40	90	3.8	460	21	AAV25456
41	90	3.8	481	21	AAV25455
42	90	3.8	522	22	AAV95513
43	90	3.8	564	21	AAV10457
44	90	3.8	564	21	AAV81952
45	90	3.8	600	22	AAV75092

ALIGNMENTS

RESULT 1

AAV85270	ID	AAV85270	standard; Protein; 442 AA.
XX	AC	AAV85270;	
XX	DT	29-JUN-2000	(first entry)
XX	DE	BASB034	amino acid sequence #3.
XX	KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging; vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia; sinusitis; nosocomial infection; invasive disease; chronic otitis media; hearing loss; antibacterial drug.	
XX	OS	Moraxella catarrhalis.	
XX	PN	WO200015802-A1.	
XX	PD	23-MAR-2000.	
XX	PF	14-SEP-1999;	99WO-EP06781.
XX	PR	14-SEP-1998;	98GB-0020002.
XX	PA	(SMIK)	SMITHKLINE BEECHAM BIOLOGICALS.
XX	PI	Ruelle J;	
XX	DR	WPI: 2000-271440/23.	
XX	DR	N-PSDB; AAA10702.	
XX	PT	Novel BASB034 polynucleotides and polypeptides from Moraxella catarrhalis used to prepare vaccines against bacterial infections	

XX PS Claim 3; Page 68; 106pp; English.

XX CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

XX CC strain Mc2913. The invention relates to BASB034 polypeptides from

XX CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034

XX CC polynucleotides and polypeptides may be employed as research reagents and

XX CC material for the discovery of treatments and diagnostics for diseases,

XX CC particularly human diseases. They are particularly used to diagnose and

XX CC treat M. catarrhalis infections. They can be used for diagnosis of

XX CC disease, staging of disease, or determining response of an infectious

XX CC organism to drugs. The polynucleotides may be used as a source for

XX CC hybridization probes, and for screening of genetic mutations, serotype,

XX CC organism or strain identification, identification of mutations in BASB034

XX CC sequences, and as components of arrays which are useful for diagnostic

XX CC and prognostic purposes. The polypeptides can be used to produce

XX CC antibodies. The polypeptides can also be used in vaccine formulations,

XX CC and to identify agonists and antagonists. The polypeptides, antibodies,

XX CC agonists and antagonists (which are bacteriostatic) are used for the

XX CC treatment and prevention of diseases such as otitis media in infants and

XX CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

XX CC invasive diseases, and chronic otitis media with hearing loss. The

XX CC polypeptides, agonists and antagonists are also used for screening of

XX CC antibacterial drugs. The BASB034 products of the invention can be used

XX CC screen for new antibacterial compounds that may target resistant

XX CC bacteria.

XX SQ Sequence 442 AA;

Query Match 100.0%; Score 2360; DB 21; Length 442;

Best Local Similarity 100.0%; Pred. No. 1.3e-218;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTSLSCFAILAIQAKAVPNPFAFVDEVSRSENDLGDQNELPIDVQSATQSA 60

DB 1 mkvslstltsilscfaiilqakavpnvpafvdevrsendlgdnelpidvqsatqsa 60

QY 61 STDATANPIDEHEPELYTTALENKTMLINCSALNQDLMRLACVDTLVHGETPAVITKRSI 120

DB 61 stdatanpidehepeilyttalenktmlincsalnqdimrlacydtlvhgetpaviktksi 120

QY 121 RUDETIWOTIKGKPOVVYQETTPDIFLMGNEKGMTKKDAKOLEYAAKQFTPLSLSFDLD 180

DB 121 rldetiwtikgkpoqvvyqettdpiflmgnekgmtkkdakqleyaakqftplsfsfld 180

QY 181 RNTPTLWSSRPHNPMYVLPFMHGKPNRSPNTPSHEARQFTNEFRAPELKFOVSVKKA 240

DB 181 rntptlwsrphnmpvlpfmgkpnrsntpshearqftnefrapelkfqvsvkka 240

QY 241 ABDLWGTSDDLWFGYTQSHWQIFNGKNSRPPRVHDYQPEIFLTPQVYSDLPWDKVRMI 300

DB 241 aedlwgtssdlwfgytqshwqifngknsrprvhdypelfltpqvysdlpwwdkvrmi 300

QY 301 GMGAVHHNGESAKLSRSWNRNRYLMAGMEWKNLTVMPTWRIWGRIFKEGSGSQDDNPDI 360

DB 301 gmgaavhngesaklsrswnrnylmagmeknltvmptwriwgrifkegsgsqddnpdild 360

QY 361 YGYGDRFLYOLENKSNTSGVRYNPRSGKALQDLYVPLGKISGYFQIFQYGSGL 420

DB 361 ygygdrfilyolenksntsgvrynprrsgkalqldlyvplgkisyfyfqqygsi 420

QY 421 IDYNHEATSFYGLMLNDWMGL 442

DB 421 idynheatsfygmlndwmg 442

RESULT 2

AY85268

AY85268 standard; Protein: 442 AA.

AY85268;

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #1.

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;

XX vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;

XX sinusitis; nosocomial infection; invasive disease; chronic otitis media;

XX hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

XX WO200015802-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

XX 14-SEP-1998; 98GB-0020002.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Ruelle J;

XX WPI: 2000-271440/23.

XX N-PSDB; AAA10700.

XX Novel BASB034 polynucleotides and polypeptides from Moraxella

XX catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Fig 2; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

XX strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides

XX from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The

XX BASB034 polynucleotides and polypeptides may be employed as research

XX reagents and material for the discovery of treatments and diagnostics for

XX diseases, particularly human diseases. They are particularly used to

XX diagnose and treat M. catarrhalis infections. They can be used for

XX diagnosis of disease, staging of disease, or determining response of an

XX infectious organism to drugs. The polynucleotides may be used as a source

XX for hybridization probes, and for screening of genetic mutations,

XX serotype, organism or strain identification, identification of mutations

XX in BASB034 sequences, and as components of arrays which are useful for

XX diagnostic and prognostic purposes. The polypeptides can be used to

XX produce antibodies. The polypeptides can also be used in vaccine

XX formulations, and to identify agonists and antagonists. The polypeptides,

XX antibodies, agonists and antagonists (which are bacteriostatic) are used

XX for the treatment and prevention of diseases such as otitis media in

XX infants and children, pneumonia in elderlies, sinusitis, nosocomial

XX infections and invasive diseases, and chronic otitis media with hearing

XX loss. The polypeptides, agonists and antagonists are also used for

XX screening of antibacterial drugs. The BASB034 products of the invention

XX can be used screen for new antibacterial compounds that may target

XX resistant bacteria.

XX SQ Sequence 442 AA;

Query Match 99.7%; Score 2352; DB 21; Length 442;

Best Local Similarity 99.3%; Pred. No. 7.6e-218;

Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTSLSCFAILAIQAKAVPNPFAFVDEVSRSENDLGDQNELPIDVQSATQSA 60

DB 1 mkvslstltsilscfaiilqakavpnvpafvdevrsendlgdnelpidvqsatqsa 60

QY 61 STDATANPIDEHEPELYTTALENKTMLINCSALNQDLMRLACVDTLVHGETPAVITKRSI 120

DB 61 stdatanpidehepeilyttalenktmlincsalnqdimrlacydtlvhgetpaviktksi 120

QY 121 RUDETIWOTIKGKPOVVYQETTPDIFLMGNEKGMTKKDAKOLEYAAKQFTPLSLSFDLD 180

DB 121 rldetiwtikgkpoqvvyqettdpiflmgnekgmtkkdakqleyaakqftplsfsfld 180

PR 14-SEP-1998; 98GB-0020002.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA Ruelle J;
 PI
 XX WPI; 2000-271440/23.
 DR N-PSDB; AAA10701.
 XX
 XX Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 XX
 XX Claim 3; Page 67; 106pp; English.
 XX
 CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which can be used for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX
 XX Sequence 442 AA;
 SQ

Query Match 98.7%; Score 2330; DB 21; Length 442;
 Best Local Similarity 98.9%; Pred. No. 1e-215;
 Matches 437; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MKVSLSTLTSLSCFAIIAQAKAVPNVAFVDEVSENDLQDNELPDIQVQSATOSA 60
 Db 1 mkvslstlslscfaiiaqqagavnpvafvdevskndlgqdnelllgvqatqsa 60

Qy 61 STDTPANPLDEHEPELYTTALENKTMLINCSALNQDIMRLACYDTLTVHGETPAVTKRSI 120
 Db 61 stdtanpldehepeilyttalenktmlincsalnqdimrlacydtlthgetpavtktrsi 120

Qy 121 RLDEITWOTIKGKPOVYQETDPIFLMGNEKGMTTKDAKOLEVAANKQFTPLSLSFOLD 180
 Db 121 rldetiwtikgkpvoyqetdtpiflmgnekgmttkdakqleyaakqftplslsfold 180

Qy 181 RNNTPLWSRRPHNMYVLPFIEMHGKPNRSPNTPSHEARQFTPNRAPELKQVSVKVA 240
 Db 181 rnntplwsrrphnmyvlpfihmgkpnrsptpshearqftpnrapelkfqsvkvka 240

Qy 241 AEDLWGTSDSLWFGYQTQSHQWQIFNGKNSRFRVHDYQPEIFLTPQVYSDLPWDGKVRMI 300
 Db 241 aedlwgtstdslwfgytqshqwwqifngknsrfrvhdypqeilftptqvysdlpwdgkvrm 300

Qy 301 GMGAVHHNSGESAKLSRSWNRAYLMAGMEWKNLTVMYPRIGWRIKFGSGSQPDNDPILD 360
 Db 301 gmgavhnsngesaklsrswnraylmagmewknltvmprigwrikgsgsqpdndpild 360

Qy 361 YYGCDVRFYQLENKNSISGTVRNPNRSGKALQDLYVPLGKIGISGYFOIFQGYGQSL 420
 Db 361 yygcdvrflyqlenknsgtvrnprnsgkalqldvvyplgkigisgyfoifqgygqsl 420

Qy 421 IDYNHEATSFVGVLMLNDMMGL 442
 Db 421 idynheatsfvgvlmlndwmg1 442

RESULT 5
 AAY75156
 ID AAY75156 standard; Protein: 370 AA.
 XX AC AAY75156;
 XX DT 21-MAR-2000 (first entry)
 XX DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX OS Neisseria gonorrhoeae.
 PN WO9957280-A2.
 XX PD 11-NOV-1999.
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AAZ53918.
 XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX Claim 2; Page 903; 1453pp; English.
 XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX Sequence 370 AA;
 SQ

Query Match 34.8%; Score 822; DB 21; Length 370;
 Best Local Similarity 44.7%; Pred. No. 1.4e-70;
 Matches 163; Conservative 68; Mismatches 110; Indels 24; Gaps 8;

Qy 87 INCSALNODIMRLACYDTLTVHGETPAVI-----KTKRSIRLDETIWQTI-KGKPOVYQTE 140

```
Db 21 lqcaaltndvtrlacydrifaalpsagqegqskavlnltetvrrssldkgeavivvek 80
QY 141 TTPDIFLMGNEKGMLTKKDAKQLEYAAKQFTPLSLGFDLDRNN-TPLWSSRRPHNPMYVLP 199
Db 81 ggda-----padsagetadlytplsmydlkndrlgllgvrehmpmyimp 127
QY 200 IFMHGKPNRSPNTPSHEAR-QFTPNFRAPELKFOVSVKAAEDLWGTDSDLWFGYTOQ 258
Db 128 fwnnspnyapsptgtttvqekfgqkaetklqvsfkkskaefktradiwfgytqr 187
QY 259 SHWQIEN-GENSRPFRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMAGVHHSNGESAKLSR 317
Db 188 sdwqinyqrksapfrntdykpeifitqpvkadlpfggrlrmigafvfhqsgsqrsper 247
QY 318 SNRRLVLMAGMEKMLTVMRPRIWGRIFKEGSGSQPDNDPILDYGYGVDFRYLYOLENKS 377
Db 248 swnrlyamagmewgkltvprvvrfafdq-sgdk-ndnpladiadymgygdvkklygrindrq 305
QY 378 NISGTVRYNPRSGKALQLDYVYPLGKIGSYGYFQIFQGYGQSLIDYNHEATSFVGGLMLN 437
Db 306 nvysvlyrnpktygaaleaaytftpikgklkvrvrgfhgygeslidyhkhqngigiglmfn 365
QY 438 DWMGL 442
Db 366 dwdgi 370

RESULT 6
AAY75157
ID AAY75157 standard; Protein; 370 AA.
AC AAY75157;
DT 21-MAR-2000 (first entry)
XX Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX Neisseria meningitidis.
OS Neisseria meningitidis.
XX WO9957280-A2.
XX 11-NOV-1999.
XX 30-APR-1999; 99WO-US09346.
XX 01-MAY-1998; 98US-0083758.
XX 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX WPI; 2000-062150/05.
DR N-PSDB; AA253919.
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
PS Claim 2; Page 904; 1453pp; English.
```

```
XX CC AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisseria bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX CC Sequence 370 AA;
QY 87 INCSALNQDLMRLACYDTLVHGETPAVI-----KTKRSIRLDETINQTI-KGKPOVYQOE 140
Db 21 lqcaaltndvtrlacydrifaalpsagqegqskavlnltetvrrssldkgeavivvek 80
QY 141 TTPDIFLMGNEKGMLTKKDAKQLEYAAKQFTPLSLGFDLDRNN-TPLWSSRRPHNPMYVLP 199
Db 81 ggda-----padsagetadlytplsmydlkndrlgllgvrehmpmyimp 127
QY 200 IFMHGKPNRSPNTPSHEAR-QFTPNFRAPELKFOVSVKAAEDLWGTDSDLWFGYTOQ 258
Db 128 lwnnspnyapsptgtttvqekfgqkaetklqvsfkkskaefktradiwfgytqr 187
QY 259 SHWQIEN-GENSRPFRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMAGVHHSNGESAKLSR 317
Db 188 sdwqinyqrksapfrntdykpeifitqpvkadlpfggrlrmigafvfhqsgsqrsper 247
QY 318 SNRRLVLMAGMEKMLTVMRPRIWGRIFKEGSGSQPDNDPILDYGYGVDFRYLYOLENKS 377
Db 248 swnrlyamagmewgkltvprvvrfafdq-sgdk-ndnpladiadymgygdvkklygrindrq 305
QY 378 NISGTVRYNPRSGKALQLDYVYPLGKIGSYGYFQIFQGYGQSLIDYNHEATSFVGGLMLN 437
Db 306 nvysvlyrnpktygaaleaaytftpikgklkvrvrgfhgygeslidyhkhqngigiglmfn 365
QY 438 DWMGL 442
Db 366 dwdgi 370

RESULT 7
AAY75158
ID AAY75158 standard; Protein; 370 AA.
AC AAY75158;
XX 21-MAR-2000 (first entry)
XX Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX Neisseria meningitidis.
OS Neisseria meningitidis.
XX WO9957280-A2.
XX 11-NOV-1999.
XX 30-APR-1999; 99WO-US09346.
XX
```


Db	193	sdwqlyngqrksapfrntdykpeliftqpvkadlpfggrrlmlgagfvhqsgngqsrpear	255
Qy	318	SWRAYLMAGMEWKNLTVMPRINWGRIFKEGSGQPDNDPILDYGYGDVRFYQLLENKS	377
Db	253	swriyamagmegwkltpirvvrfafdq-sgdk-ndnpdiadymgygdvklqyrlndrq	310
Qy	378	NISGTVRYNPRSGKALQLDYVYPLCKGIGSYQFQITQGYGQSLIDYNHEATSFQVGLMLN	437
Db	311	nvysvlyrnpktygaaiaaaytpikgklgvvrgfhgvgeslidynhknqngigglmfn	370
Qy	438	DWMGL 442	
Db	371	dlqgi 375	
RESULT 10			
AAW98871			
ID	AAW98871	standard; Protein; 355 AA.	
AC	AAW98871;		
XX			
DT	31-MAR-1999	(first entry)	
XX			
DE	H. pylori	GHPO 1723 protein.	
XX			
KW	GHPO protein; Helicobacter infection; gastroduodenal disease; gastrit		
KW	peptic ulcer disease.		
OS	Helicobacter pylori.		
PN	WO9843478-A1.		
XX			
PD	08-OCT-1998.		
XX			
PF	01-APR-1998; 98WO-0506371.		
XX			
PR	29-JUL-1997; 97US-0902615.		
PR	01-APR-1997; 97US-0833457.		
PR	24-JUN-1997; 97US-0881227.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PA	(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.		
PI	Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;		
XX			
DR	WPI; 1998-542293/46.		
DR	N-PSDB; AAX14590.		
PT	New isolated Helicobacter polynucleotides - used to develop products		
PT	for the diagnosis, prevention and treatment of Helicobacter		
PT	infections and gastrointestinal diseases		
XX			
PS	Claim 8; Page 1976-1977; 2054pp; English.		
XX			
CC	This sequence represents a Helicobacter pylori GHPO protein of the		
CC	invention. The polypeptides can be used for preventing or treating		
CC	Helicobacter infections, and gastroduodenal diseases associated with		
CC	these infections, including acute, chronic, and atrophic gastritis, a		
CC	peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also		
CC	used for the production of antibodies. The products can also be used		
CC	detection and diagnosis.		
XX			
SQ	Sequence 355 AA;		
Query Match 10.4%; Score 246.5; DB 19; Length 355;			
Best Local Similarity 25.1%; Pred. No. 3.7e-15;			
Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps			
Qy	157	KDKAQLEYAAKQFTPLUSLFSFLDRNNTPLWSSRPHNPMYVLPFVHMGKPNRSPNTPSHE	216
Db	69	kkvlnmmdvltqvtglbfvhsf-----tqifawyhoninp-----	102

QY 217 ARQFTPNRPAELKFOVSQVKKAAEDLWGTDSDLMFGYTOQSHQWIFNGKNSRPRFRVHD 276
 Db 103 ---yrnef-----kfisfrvpvrfrhllwtktlylaytqtdwfdqyndpqsapmrmmn 154
 QY 277 YOPEIFLTOPVYSDLPWDCGV---RMIGMGAVHHSNG-ESAKLSRSWNRAYLMAGMEWKN 332
 Db 155 fmpellyvyypf-nfkpfggkignfseiwghisngvggacypfnk-----egnpnq 209
 QY 333 LVMMPRI-----WGRIFKSGSQSP-----DDNDP 357
 Db 210 fggqpvivkdyngqkdvrgwgrsvsagrpvrfrlvwkggikimvayvpyvpydqsnpn 269
 QY 358 ILDYGYGVDVRELY-----OLENKSNTSGTVRYNPRSGKALQLDYVYPLGKIGSYGF 410
 Db 270 lidymgynakidyrgrhhfelqlydiftqwyrd--rwghafriygyrinpfvgiya 327
 QY 411 QIFOGYGOSLIDYNEHATSFGVGLMLN 437
 Db 328 qwfdyngdglyeydvfsnrigvgirln 354

RESULT 11
 ID AAY10960 standard; Protein; 356 AA.
 XX AAY10960;
 DT 08-JUN-1999 (first entry)
 DE H. pylori ORF 07ap80601_5083193_f3_8 cell envelope protein.
 KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
 KW secreted protein; cellular protein.
 OS Helicobacter pylori.
 XX WO981823-A1.
 PD 07-MAY-1998.
 PF 28-OCT-1997; 97WO-US19575.
 XX 14-JUL-1997; 97US-0891928.
 PR 28-OCT-1996; 96US-0739150.
 XX 06-DEC-1996; 96US-0759739.
 PA (ASTR) ASTRA AB.
 XX Alm RA, Smith D;
 XX WPI; 1998-271811/24.
 DR N-PSDB; AAX30427.
 PT Helicobacter pylori nucleic acids and proteins - used to develop
 PT products for the detection, prevention and treatment of H. pylori
 PT infections
 XX Claims 27, 31; Page 160-161; 279pp; English.
 XX Recombinant or substantially pure preparations of H. pylori polypeptides
 CC are disclosed, together with the nucleic acids encoding them. In all,
 CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
 CC secreted proteins or other cellular proteins. Vaccines containing the
 CC nucleic acids or proteins are claimed, as are probes containing at least
 CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful
 CC for treating or reducing the risk of H. pylori infections, and the
 CC probes can be used diagnostically for detecting the presence of
 CC Helicobacter in a sample. The products are also of use in screening
 CC for compounds having the ability to interfere with the H. pylori life
 CC cycle or to inhibit H. pylori infection.
 XX Sequence 356 AA;

Query Match 10.1%; Score 239.5; DB 19; Length 356;
 Best Local Similarity 24.8%; Pred. No. 1.7e-14;
 Matches 82; Conservative 43; Mismatches 111; Indels 95; Gaps 12;
 QY 157 KDAKQLEVAQAQFTPLSLSFOLDNRNTPMLSSRPHNPWYVLPIDFMHGKPNRSPNTPSHE 216
 Db 70 kkylnmmdylgtfipfyhsf-----tpifqwyhpninp----- 103
 QY 217 ARQFTPNRPAELKFOVSQVKKAAEDLWGTDSDLMFGYTOQSHQWIFNGKNSRPRFRVHD 276
 Db 104 ---yrnef-----kfisfrvpvrfrhllwtktlylaytqtdwfdqyndpqsapmrmmn 155
 QY 277 YOPEIFLTOPVYSDLPWDCGV---RMIGMGAVHHSNG-ESAKLSRSWNRAYLMAGMEWKN 332
 Db 155 fmpellyvyypf-nfkpfggkignfseiwghisngvggacypfnk-----egnpnq 209
 QY 323 -----YLMAGMEWK---NLTVMRPRIGRIFKEGSGSQSPDD- 354
 Db 215 pviwkdyngqkdvrgwgrsvsagrpvrfrlvwkggikimvayvpyv-----pydq 266
 QY 355 -NPDIIDYGYGVDVRELY-----OLENKSNTSGTVRYNPRSGKALQLDYVYPLGKIGI 406
 Db 267 snpqlidymgynakidyrgrhhfelqlydiftqwyrd--rwghafriygyrinpfv 324
 QY 407 SGYFQIFOGYGOSLIDYNEHATSFGVGLMLN 437
 Db 325 giyaqwfngydglyeydvfsnrigvgirln 355

RESULT 12
 ID AAW20760 standard; protein; 253 AA.
 XX AAW20760;
 DT 15-JUL-1997 (first entry)
 DE H. pylori outer membrane protein, 07ap80601orf8.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;
 KW outer membrane.
 OS Helicobacter pylori.
 XX WO9640893-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; 96WO-US09122.
 PR 01-APR-1996; 96US-0630405.
 PR 07-JUN-1995; 95US-0487032.
 XX (ASTR) ASTRA AB.
 XX Berglindh OT, Smith D, Wellgaard BL;
 XX WPI; 1997-052306/05.
 DR N-PSDB; AAT68013.
 XX Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 XX Claim 56; Page 1172-1173; 1481pp; English.
 XX The present sequence is a Helicobacter pylori outer membrane protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, useful
 CC as potential H. pylori life cycle activators or inhibitors. The genomic

CC sequence of *H. pylori* (ATCC 55679) was determined from overlapping
 CC contigs generated by mechanically shearing the bacterial DNA. The
 CC sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely *H. pylori* antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from *H. pylori* by PCR amplification for recombinant polypeptide
 CC production, e.g. in *E. coli* hosts.

XX Sequence 253 AA;

Query Match 6.6%; Score 156.5; DB 18; Length 253;
 Best Local Similarity 25.6%; Pred. No. 1e-06;
 Matches 56; Conservative 26; Mismatches 72; Indels 65; Gaps 8;

QY 157 KDKAKOLEYAAKQFTPLSLSFDLDRNTPMSSRPHPNPMVLPDIFMHGKPNRSPNTPSHE 216
 DB 72 kkylnmndylgtfplfyhsf-----tpifqwyhpnlnp----- 105

QY 217 ARQFTNEFRAPELKFOVSVKVKAAEDLAGTDSDLWFGYTOQSHWOIFENGKNSRPFVHD 276
 DB 106 ---yqrnef-----kqisfrvfpvfhilwtktgtylaytqtnwfglynopdsapmrmin 157

QY 277 YOPEIFLTPVYSDLPDWGRV---RMIGMGAVHHSNG-ESAKLSRSWNRAYLMAGMEWKN 332
 DB 158 fmpeliyyvpi-nfkpfqkignfseiwgwhisngvgagcyqfn----- 204

QY 333 LTVMPRIWGRIFKEGSGS--PDDNPDIIDYGYGVDVRF 369
 DB 205 -----kegnpenqfpgqpivkdyngqkdvrw 231

RESULT 13

AAW70991 ID AAW70991 standard; Protein; 1686 AA.

XX AC AAW70991;

XX 19-OCT-1998 (first entry)

XX Human class II P13 kinase-C2alpha.

XX Human; class II phosphoinositide lipid kinase; P13 kinase;
 KW PIK3-C2alpha; class II; resistance; Wortmannin; LY294002.

XX Homo sapiens.

XX W09832864-A2.

XX 30-JUL-1998.

XX 27-JAN-1998; 98WO-GB00244.

XX 28-JAN-1997; 97GB-0001652.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Domin J, Waterfield MD;

XX WPI; 1998-427960/36.

XX N-PSDB; AAV42920.

XX New nucleic acid encoding phosphoinositol kinase 3-C2 alpha or its
 PT fragments - useful for, e.g. treatment of tumour cells where
 PT phenotype is associated with expression of kinase

XX Claim 3; Fig 1; 52pp; English.

XX The present sequence represents a human class II phosphoinositide lipid
 CC (P13) kinase designated PIK3-C2alpha. It is characterised as a class II

CC kinase due to the presence of a conserved C2 domain found in murine and
 CC Drosophila class II P13 kinases, its apparent lack of a p85 binding site
 CC and a substrate affinity to inositol lipids PtdIns and PtdIns(4)P. The
 CC protein has resistance to P13 kinase inhibitors Wortmannin and LY294002.
 CC Antibodies against the protein (optionally humanised), are used to
 CC identify class II P13 kinases. Antisense sequences, antibodies or
 CC dominant negative mutants of the P13-C2alpha protein, are useful in human
 CC or veterinary medicine to block class II kinases. They can be used to
 CC treat tumour cells where the phenotype is associated with expression of
 CC P13-C2alpha protein.

XX Sequence 1686 AA;

Query Match 4.8%; Score 114; DB 19; Length 1686;
 Best Local Similarity 19.0%; Pred. No. 0.25;
 Matches 85; Conservative 68; Mismatches 154; Indels 140; Gaps 21;

QY 45 QDNELPIDV-----QSATOSASTDTANPLDE-----HEPELYTTALENKTMLIN----- 88
 DB 522 eddetpvdlnkhlygiekpcckeamtrhpveellidsyhnqvelalqlenqhravdqvikav 581

QY 89 ---CSALNQDIMEPLACYDILVHGETPAVVKTKRSIRLDETITWOTIKGKPOVVYQETDPI 145
 DB 582 rkicsald-gvetial-----tesvkkikravnlpr-----ktadv 618

QY 146 FLMGNEKGLTKKDAKQLEYAAKQFTPLSLSFDLDRNTPMSSRPHPNPMV-----LP 199
 DB 619 slfge-----dtsrst-rgalnpenpvgvslngltaa 651

QY 200 IF-----MHGKPNRSPNTPSHEARQFTNEFRAPELKFOVSVKVKAAEDL---WGTDSDLW 252
 DB 652 iydllrlhansgrsptdcagsskvskeawttteqlgtif---aahgissnvwvnyeky 707

QY 253 FGVTQOSHWOIFNGKN-SRPF---RVHDIQPEIFLQ-----PV-YSLPDWGRVMI 300
 DB 708 ylicslsh-----ngklfkipqskkvgtynkffylkwdellilfpiglsqlesvhlht 763

QY 301 GMGAVHHSNGESAKLSRSWN-----RAYLMAGMEWKNLTVMPRIWGRIFK 345
 DB 764 lfiglnqsgsspsdnkqrkgpealgkvalplcdfrflltcg-----tkllylv----- 812

QY 346 EGSQSPDDNPDIIDYGYGVDVRFYQLENKSNISGTVRYNPRSGKALQLDVYVPLGKG 405
 DB 813 --tsstnsvpgtvttkgyvmerivlgvdfpsafdiitytpqvdrrsilqghnletlend 870

QY 406 ISGVFQIFQYGGOSLIDYNHEATSFGV 432

DB 871 ikg-----klidilhkdsllgl 887

RESULT 14

AAW70420

ID AAW70420 standard; protein; 800 AA.

XX AC AAW70420;

XX 20-JAN-1991 (first entry)

XX Sequence encoded by cellulase gene derived from *Bacillus* sp. No. 1139.

XX Enzyme; cellotriose; cellotetrose; hydrolysis.

XX *Bacillus* sp. No. 1139.

XX Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..800

XX JP62232386-A.

XX 12-OCT-1987.

XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:35 ; Search time 44.45 Seconds
(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 2363
Sequence: 1 MKVSLSTLTLILPCFAILA.....YNHEATSFVGLMLNDWMGL 442
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	815	34.5	382	2	H81195 phospholipase A1,
2	815	34.5	409	2	H81831 probable phospholip
3	373	15.8	286	2	B36971 outer membrane pho
4	366	15.5	329	2	D81279 phospholipase A1 (
5	358	15.2	289	2	A36971 outer membrane pho
6	357	15.1	289	1	PSECA1 phospholipase A1 (
7	357	15.1	289	2	E86069 outer membrane pho
8	321.5	13.6	289	2	C36971 outer membrane pho
9	246.5	10.4	355	2	C64582 phospholipase A1 p
10	240.5	10.2	355	2	H71930 probable phospholi
11	111.5	4.7	602	1	TVTRR protein kinase (EC
12	108.5	4.6	800	2	A29003 cellulase (EC 3.2.
13	108.5	4.6	822	2	JT0611 cellulase (EC 3.2.
14	107.5	4.5	824	3	JC7532 cellulase (EC 3.2.
15	106.5	4.5	901	2	T20122 hypothetical prote
16	100.5	4.3	660	2	E83656 methionyl-tRNA syn
17	100.5	4.3	783	2	JC5467 cellulase (EC 3.2.
18	100.5	4.3	1658	2	T42642 phosphoinositide 3
19	99.5	4.2	797	2	T46737 X-Pro dipeptidyl-p
20	99	4.2	5005	2	F82884 hypothetical prote
21	98.5	4.2	810	2	S49744 AMP deaminase (EC
22	98.5	4.2	838	2	A66557 probable receptor
23	96	4.1	791	2	H96839 hypothetical prote
24	96	4.1	4307	2	T20721 hypothetical prote
25	95.5	4.0	719	2	A42893 penicillin-binding
26	95	4.0	765	2	T35719 chitinase - Strept
27	95	4.0	888	2	S50801 AMP deaminase homo
28	94.5	4.0	397	2	A35136 cellulase (EC 3.2.
29	94.5	4.0	601	2	T26062 hypothetical prote

30	94.5	4.0	655	2	T26061	hypothetical prote
31	94.5	4.0	1310	2	T40135	probable involueme
32	94.5	4.0	1478	2	S20117	protein kinase BCK
33	93.5	4.0	564	2	T40777	ferric reductase t
34	93.5	4.0	719	2	S28031	penicillin-binding
35	93.5	4.0	719	2	S28033	penicillin-binding
36	93.5	4.0	857	1	A41369	S-receptor kinase
37	93	3.9	1078	2	T18352	protein P120 - Myc
38	92.5	3.9	486	2	S30959	gene 14 protein -
39	92.5	3.9	719	2	S28034	penicillin-binding
40	92.5	3.9	719	2	S28032	penicillin-binding
41	92	3.9	324	2	B69521	hypothetical prote
42	92	3.9	470	2	T43675	cog-2 protein - Ca
43	91.5	3.9	796	2	JC7355	peroxisome prolife
44	91.5	3.9	1641	2	D82704	conserved hypothet
45	91	3.9	367	2	T24058	hypothetical prote

ALIGNMENTS

RESULT 1

E81195
phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: E81195
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755
A:Accession: E81195
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <TET>
A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40901.1; PID:g722
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0464

Query Match	34.5%	Score 815;	DB 2;	Length 382;
Best Local Similarity	44.7%	Pred. No. 3.5e-58;		
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;
QY 87	INCSALNQDIMRLACYDTLVHGETPAVI-----KTKRSIRLDETIWOTI-KGKPQVYVQE	140		
DB 33	LOCAALTNDVTRLACYDRIFAAQLPSSAGQEQESKAVLNLTETVRSLLDKGEAVIVVEK	92		
QY 141	TTDPIFLMGNEKMLTKDAKQLEYAAKQFTPLSLFSDLDLDRNN-TPLWSSRPINPMYVLP	199		
DB 93	GGDAL-----PADSAGETADITYTPLUSMDYDLDKDLRLGLLVGRHNPMYVLP	139		
QY 200	IFMHGPNRSPNTPSH-EAKQFTPNFRAPFLQSVKVAEADLWGTDSDLWFGYTOQ	258		
DB 140	LWYNNSPNAPGSPTRGTTVQEFQGGKRAETKLQVSFKSKIAEDLFKTRADLWFGYTQR	199		
QY 259	SHWQIFN-GKNSRPRFVHDYQPEIFITQPVYSDLPWDGKVRMIGMGAHVHNSGESAKLSR	317		
DB 200	SDWQIYNQGRKSAFFRNTDYKPEIFLTQPVKADLPFGGRLRLGAGFVHQSGQSRPESR	259		
QY 318	SNWRYLMAGMEKMLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGVDRFELYQLENKS	377		
DB 260	SNWRYLMAGMEKMLTVIPRVWVRAFQD-SGDK-NDNPDIAVMGYGVKQLQRLNDRQ	317		
QY 378	NTSGTVRYNPRSGKALQLDYVYPLGKIGSGYQIFQGYGQSILIDYNHEATSFQVGLMLN	437		
DB 318	NYSVLRYNPKTGYGAIEAAYTPKIGKLGKGVVRGFHGYGESLIDYNHKONGIGLMLFN	377		
QY 438	DWML 442			

Db 378 DLDGI 382

RESULT 2

H81831

probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serogroup C):Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81831

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jajelski, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:2022556

A:Accession: H81831

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738065

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2021

Query Match 34.5%; Score 815; DB 2; Length 409;

Best Local Similarity 44.7%; Pred. No. 3.9e-58;

Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

Qy 87 INCSALNODIMRLACYDILVHGETPAVI-----KYKRIRLDETWTQRI-KGKPOVYQOE 140

Db 60 LOCAALTDNVTRLACYDRIFAQLPSSAQEGQESKAVLNTETVRSLLDKGEAVIVVEK 119

Qy 141 TTDPIFLMGNEKGMTKDAKQLEVAAKQFTPLSLFSDLDNRNN-TPLNRRPHNPWVLP 199

Db 120 GGDAL-----PADSAGETADIIYPLSLMYDLDDKDLRGLGVRHNPWVLP 166

Qy 200 IFMHGKPNRSPNTPSH-BAKQFTPNFRAPELKFQVSVKAAEDLWGTDSDLWFGYQQ 258

Db 167 LWYNSPNYAPGSPRGTTVQEKFGQKRAETKLVQVSEKIAEDLEKTRADLWFGYTOR 226

Qy 259 SHWQLEN-GKNSRPRVHDYQPEILTPQVYSDLPWDGKVRMIGMGAHVHNSGESAKLSR 317

Db 227 SDWQYNGRKSAPPNTDYKPEIFLTQPVKADLPFGGLRLMLGAGFVHQSGQSRPESR 286

Qy 318 SNRAYLMAGMEWKKLTVMPIRWIRIFKEGSGQPDNDPDIIDYGYGQVRFYOLENKS 377

Db 287 SNRIYAMAGMEWKKLTIPRVWVPAFDQ-SGDK-NDNPDADYMGYGDVKLOYRLNDRQ 344

Qy 378 NISGTVRYNPRSGKALQLDYVYPLGKISGYFTQFGYQSGSLIDYNHEATSFVGGLMLN 437

Db 345 NVYSVLRYNPKTGYGAIEAAVTFPIKGLKGVVRFGHGYGESLIDYNHKNQNGIGIGLMFN 404

Qy 438 DWGL 442

Db 405 DLDGI 409

RESULT 3

B36971

outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae

C:Species: Klebsiella pneumoniae

C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999

C:Accession: B36971; S40129

R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm J. Bacteriol. 176, 861-870, 1994

A:Title: Molecular characterization of enterobacterial plga genes encoding outer membrane phospholipase A

A:Reference number: A36971; MUID:94131966

A:Accession: B36971

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <BRO>

A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881

A:Note: authors translated the codon AAG for residue 112 as Arg

C:Genetics:

A:Gene: plda

C:Superfamily: bacterial phospholipase A1

C:Keywords: carboxylic ester hydrolase

Query Match 15.8%; Score 373; DB 2; Length 286;

Best Local Similarity 42.1%; Pred. No. 1.2e-22;

Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

Qy 226 RAPELKFQVSVKAAEDLWGTDSDLWFGYQQSHWQIFNGKNSRPRVHDYQPEIFITQ 285

Db 80 RKDEVKFQSLAFPLPILGILGDSLLGASYTQKSWQLSNKSESAPFRETYEPQLFLGF 139

Qy 286 PV-YSDLPWGDKVRMIGMGAHVHNSGESAKLSRSNRAYLMAGMEWKKLTVMPIRWIRIF 344

Db 140 ATDYQFAGW--TLRDIEMGYNHDSNGRSDPTSRNRLYARLMAQNGNLVEVKPW---Y 194

Qy 345 KEGSGSQPDNDPDIIDYGYGQVRFYOLENKSNTSGTVRYNPRSGKALQLDYVYPLGK 404

Db 195 VVGS---TDDNPDIYMGYVRLKVGYL-GEAIISSAQGYNNWNTGYGGAELGVSYPITK 250

Qy 405 GISGYFTQFGYQSGSLIDYNHEATSFVGGLMLND 438

Db 251 HVRAYTQIYSGYGESLIDYNFNQTRGVGGLMLND 284

RESULT 4

D81279

phospholipase A1 (EC 3.1.1.32) Cj1351 [imported] - Campylobacter jejuni (strain NCTC C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000

C:Accession: D81279

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals A:Reference number: A81250; MUID:20150912

A:Accession: D81279

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-329 <PAR>

A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73778.1; PID:g696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: plda; Cj1351

C:Keywords: carboxylic ester hydrolase

Query Match 15.5%; Score 366; DB 2; Length 329;

Best Local Similarity 33.2%; Pred. No. 5.7e-22;

Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

Qy 130 IKGRPQVYQETTDPIFLMGNEKGMTKDAKQLEVAAKQFTPLSLFSDLDNRNT--PLW 187

Db 43 LKNSVLSIQEQNNSSQATQNSITIKKEKQ-----DFSRLALANYLGENSESFPLG 96

Qy 188 SSRPHNPVLPPIFMHGPKNRSPNTPSHEAKQFTNEFRAPELKFQVSVKAAEDLNGT 247

Db 97 IS-SYKMYFLP-FAYSFNSLGVNNKSEA-----RFQLSVKKRLFENLGL 141

Qy 248 DSDLWFGYQQSHWQIFNGKNSRPRVHDYQPEIFLTQPVY-SDLPWGKVRMIGMGAHV 306

Db 142 DEKYIATQTSWQIY--EHSSPFRNTYQPEFIDLPYLKDYEFFNNLR---VGILH 196

Qy 307 HNSGESAK--LSRSNRAYLMAGMEWKKLTVMPIRWIRIFKEGSGSQPDNDPDIIDYGY 364

Db 197 ESNKGKDENLQSRWNRIYVSTAILYNKFLFVPLWYRI---PENKDDDNPAIILHMG 253

Qy 365 GDVRFYOLENKSNTSGTVRYNPR--SCKGALQLDYVYPL-GKISGYFTQFGYQSGSLI 421

Db 254 FDNLAY-LGDDYFINLMLRNKLFHNNKGAIQVDIGYDIFNNGIYWLQYFNGYGESLI 312

[illegible]


```
Query Match 10.2%; Score 240.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 8.9e-12;
Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KDKAKOLEYAAKQFTLSLSDLDNRNTPWSSRPHNPMYVLPFIMHGKPNRSPNTPPSHE 216
Db 69 KLYLNMDYLGTYFLPYHSF-----TPFQWYHPNINP----- 102

QY 217 AKQFTNEFRAPLEKQVSVKYKAAEDLWGTDSLDLWFGYTOQSHWQIFNGKNSRPRVRVD 276
Db 103 ---YORNEP---KQISFRVPEVRHILTKGTLYLAYTQTNWFOYNDPQSPAPRMIN 154

QY 277 YQEIFLTPQVSDLPWDQKV-----RMIGCAVHHNG-ESAKLSRSWNR----- 321
Db 155 FMPELIYVPI-NFKPFGKIGNFSEIWIQWHSISGVGAQCYQPFNKEGPNQFPQG 213

QY 322 -----AYLMAG-----MEWK--NLTVMPRIWGRIFKEGSGSQPDD- 354
Db 214 PVIVKDYNGQKVRWGCRSVSAGNALCFVLVWEKGLKIMWAYWPYV-----PYDQ 265

QY 355 -NPDIILDYGYGDVRFY-----OLENKSNIISGTVRYNPRSGKGALQLDYVPLGKI 406
Db 266 SNQOLIDYMGYGNKADYRGRHHFELQLYDFTQWRVD--RWHGAFRLGYTYRINPFV 323

QY 407 SGYFQIQGYGQSLIDYNHEATSFVGGLMLN 437
Db 324 GIYAQWFGYGDGLYEYDFVFSNRIGVGIRLN 354

RESULT 11
TVTRR
N:Alternates: kinase (EC 2.7.1.37) raf - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: B26126
R:ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
Mol. Cell. Biol. 7, 1226-1232, 1987
A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein.
A:Reference number: A26126; MUID:87172791
A:Accession: B26126
A:Molecule type: mRNA
A:Residues: 1-602 <LSH>
A:Cross-references: GB:M15428; NID:G206546; PIDN:AAA42002.1; PID:G206547
C:Genetics:
A:Gene: raf
C:Superfamily: rat protein kinase raf; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonine
F:301-567/Domain: protein kinase homology <KIN>
F:309-317/Region: protein kinase ATP-binding motif
F:329/Active site: Lys #status predicted
F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted

Query Match 4.7%; Score 111.5; DB 1; Length 602;
Best Local Similarity 19.3%; Pred. No. 0.53;
Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;

QY 23 QAQAVPNVAFVDEVRSENDLGQDNELPIDVQSATQSASTDTANPLDEHEPELYTTALEN 82
Db 81 ELEKLNQVNVLTPEKNKELETAQDNLGI-----QSQFTRAKEELEAEKRLDIRT---N 131

QY 83 KTLNCSALNODIMRLACYDYLHGHGTAVIKTKRSIRLDETIWQ--TIKGRPOVYQOE 140
Db 132 ERUSQVEYLTDVVKRL--NEKLKESNT---TKGELQLQLDELQASDVTVKYREKRLQOE 186

QY 141 -----TTDPIFLMGNKSG-----MLTKKDAKOLEYAAKQFTPLS 174
Db 187 KELLHNQSWNLNTELYTKTDLELLAGREKGNELTELEKCTLENKKEADAIKSHSESAPSA 246

QY 175 LSFELDRNNTPLSSRPHNPMYVLPFIMHGKPNRSPNTPPSHEAKQFTPNFRAPLEKQV 234
Db 247 LS-SSPNLSPTGWSQPKTP-----VPAQRERAPCGSGTQEKNIKIRPRQORDSSYYWEI 298
```

```
QY 235 SVKVKAAEDLWGTDSLDLWFGYTOQSHW-----QIFNGKNSRPRVRVDYQPEI----- 281
Db 299 EASEVMLSTRIGSGS---EGTVYKWKWHDGVAVKILKVVDDPTPEQLQAFRNEAVLKRKR 355

QY 282 -----FLTQ-----PVYSDL--PWGKGVKRM-----GMCAVHH 307
Db 356 HVNILLFMGYMTKDLAIVTQMCESGLSKHLHVQETKTFQMFQFQIIDIARTAGMDYLLHA 415

QY 308 SNGESAKLSRSWNRAYLMAGMEWK---NLTVMPRIWGRIFKEGSGSQ----- 351
Db 416 KNIHRDMKS--NNIFLHEGLTVKIGDFGLATVKSRL-----SGSQVQEQPTGSLVLM 466

QY 352 -----PDNPDIL--DYGYGDVRFYOLENKSNIISGTVRYNPRSGKGALQLDYVYPL 402
Db 467 APEVIRMOONNPFSFQSDVSYGIV--LYEL-----MTGELPYSHINNRDQI---IFMV 515

QY 403 GK 405
Db 516 GRG 518

RESULT 12
A29003
cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2000
C:Accession: A29003
R:Fukumori, F.; Kudo, T.; Narahashi, Y.; Horikoshi, K.
J. Gen. Microbiol. 132, 2329-2335, 1986
A:Title: Molecular cloning and nucleotide sequence of the alkaline cellulase gene fro
A:Reference number: A29003; MUID:87085443
A:Accession: A29003
A:Molecule type: DNA
A:Residues: 1-800 <PUK>
A:Cross-references: GB:D00066; GB:N00066; NID:G216223; PIDN:BAA00045.1; PID:G216224
A:Experimental source: strain 1139
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
A:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 108.5; DB 2; Length 800;
Best Local Similarity 19.9%; Pred. No. 1.4;
Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

QY 8 LTLNILPCFAILAIQQAQAVPNVAFVDEVRSEND---LGQDNELPIDVQSATQSASTDT 64
Db 10 LISSILILVLLSL-----FPTALAEQVTRDNFKHLLGNDNVKRPSEAGALQLEVDG 64

QY 65 ANPL-DEHEPELY-----TTALENKTMLNCSAL-----NQDIMRLACY-DTLVHGET 110
Db 65 QMTLVQHGKEIKLRGSMTHGLQWFPPEILNDNAYKALANDWESNMIRLAVGVGNGVASN 124

QY 111 PAVIKYKRSIRLDETI-----NQI-----TIKGRPOV 137
Db 125 PELIKSRVTKGIDLATENDMYIVDWHVHAPGDPVPVAGAEDEFRRDIAALYPNPHII 184

QY 138 YQETTPD-----IFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSDLDNRN---NTPL 186
Db 185 YELANEPSSNNNGAGIPNNEEGWNAVK-----EYADPIVEMLRDSGNADNIIIVGSPN 239

QY 187 WSSRP-----HNPMYVLPFIMHGKPNRS--PNTPSHEAKQFTPNFRAPLEK 232
Db 240 WSORPDLAADNPIDDDHTMYTVHFYTGSHAASSTESPPETPNSEKGNVMSNTRYA---L 295

QY 233 QVSVKYKAAEDLWGT-----DSDLWFGYTOQ-----SHWQIFNGKNS-----R 270
Db 296 ENGVAVFATE--WGTQANGDGGPYDEADWIEFLNENNISWANWLSLTN-KNEVSGAFT 352
```

Qy 271 PFRVH-----DYQPEIFLQPVYSDLPW--- 293
|| :
Db 353 PFLGKSNATSLDPCGDQVWPPEELSLSGEYVRARIKGVNVEP-----IDRTKYTKVLWDFN 409

Qy 294 DGKVRMIGAVHHSNGES-----AKLSRSWNRAYLMAGWEKNTLTMVPRIW 340
|| :
Db 410 DGTGKGFV-----NGDSPVEDVVIENEGALKLSGLDASNDVSEGNWYANARLSADGW 463

Qy 341 GR 342
|| :
Db 464 GK 465

RESULT 13
JT0611
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
C:Accession: JT0611
R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
Biosci. Biotechnol. Biochem. 56, 872-877, 1992
A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic bacterium
A:Reference number: JT0611; MUID:92305459
A:Accession: JT0611
A:Molecule type: DNA
A:Residues: 1-822 <SUM>
A:Cross-references: GB:M84963; NID:g289264; PIDN:AAA73189.1; PID:g289266
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel-
A:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal rep
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F:585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 108.5; DB 2; Length 822;
Best Local Similarity 19.9%; Pred. No. 1.5;
Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

Qy 8 LTLSLPCFAILATQQAQVNPVAFVDEVRSEND---LGQDNELPIDVQASQTASTDT 64
|| :
Db 10 LISSILILVLLSL-----FPTLAAEGNTREDNFKHLGNDNVKRPSEAGALQLEVDG 64

Qy 65 ANPL-DEHEPELY-----TTALENTMLINGCAL-----NQDIMRLACY-DTLVHGET 110
|| :
Db 65 QMTLVDQHGKIQLRGMSTHGLWFPPEILNDNAYKALANDNESNMIRLAMYVGENGYASN 124

Qy 111 PAVIKTRSIIRLDETI-----WQT-----IKGKQVY 137
|| :
Db 125 PELIKSRVIGIDLAIENDMYVVDVHVAHAPGDPDPVYAGAEDEFRIAALYPNNPHI 184

Qy 138 YQETTD-----IFLMGNEKGMTKKDAKQLEYAAKQFTPLSLDFLDNRN-----NTPL 186
|| :
Db 185 YELANEPSSNNGGAGIPNNEEGNAVK-----EYADPIVEMLRKSGNADNIIIVGSPN 239

Qy 187 WSSRP-----HNPWVLPF--MHGKPNRS--NTPSHEAKQFTPNFEAPELKF 232
|| :
Db 240 WSRQPLADNPIDDHHTWYVHTYGTGSHAASTESYPTPNSEGNVMSNTRYA-----L 295

Qy 233 QVSVKVAADLWGT-----DSDLWFGYTOQ-----SHWQIFNGKNS-----R 270
|| :
Db 296 ENGVAVFATE--WGTSQANGDGGPYDEADVWIEFLNENNISWANNSLTN-KNEVSGAFT 352

Qy 271 PFRVH-----DYQPEIFLQPVYSDLPW--- 293
|| :
Db 353 PFLGKSNATSLDPCGDQVWPPEELSLSGEYVRARIKGVNVEP-----IDRTKYTKVLWDFN 409

Qy 294 DGKVRMIGAVHHSNGES-----AKLSRSWNRAYLMAGWEKNTLTMVPRIW 340
|| :
Db 410 DGTGKGFV-----NGDSPVEDVVIENEGALKLSGLDASNDVSEGNWYANARLSADGW 463

Qy 341 GR 342
|| :
Db 464 GK 465

RESULT 14
JC7532
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-S237)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp. (strain KSM-S237)
C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7532; PC7107
R:Hakamada, Y.; Hatada, Y.; Koike, K.; Yoshimatsu, T.; Kawai, S.; Kobayashi, T.; Ito, Biosci. Biotechnol. Biochem. 64, 2281-2289, 2000
A:Title: Deduced amino acid sequence and possible catalytic residues of a thermostabl
A:Reference number: JC7532; MUID:21036886
A:Accession: JC7532
A:Molecule type: DNA
A:Residues: 1-824 <HAK>
A:Cross-references: DDBJ:AB018420
A:Experimental source: strain KSM-S237
A:Accession: PC7107
A:Molecule type: protein
A:Residues: 31-50 <HA2>
C:Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent
C:Genetics:
A:Gene: Egl-237
C:Keywords: hydrolase; glycosidase

Query Match 4.5%; Score 107.5; DB 3; Length 824;
Best Local Similarity 19.7%; Pred. No. 1.8;
Matches 91; Conservative 56; Mismatches 147; Indels 167; Gaps 27;

Qy 30 PVAFVDEVRSEND---LGQDNELPIDVQASQTASTDTANPL-DEHEPELY-----TT 78
|| :
Db 25 PAALAAEGNTREDNFKHLGNDNVKRPSEAGALQLEVDGQMTLVDQHGKIQLRGMSTH 84

Qy 79 ALENKTMILINGCAL-----NQDIMRLACYDTLHVGE-----TPAVIKTK-----RSI 120
|| :
Db 85 GLWFPPEILNDNAYKALSNWDNSNIRLAMY-----VGENGYATNPPELIKQKVIDGIELAI 140

Qy 121 RLDETI---WQT-----IKGKQVYVQETDP-----IF 146
|| :
Db 141 ENDMYVVDVHVAHAPGDPDPVYAGAKDFREIAALYPNNPHIYELANEPSSNNGGAG 200

Qy 147 LMGNEKGMTKKDAKQLEYAAKQFTPLSLDFLDNRN-----NTPLWSSRP-----H 192
|| :
Db 201 IPNNEEGKAVK-----EYADPIVEMLRKSGNADNIIIVGSPNWSQRPDLAANDPIDDH 255

Qy 193 NPMVVLPIFMHGKPNRSNTPSHEAKQFTPNFEAPELK-----FQVSVKVAADLWGT 247
|| :
Db 256 HTWTVVHTYGTGSHAASTESYSE-----TPNSERGNVMSNTRYALENGVAVFATE--WGT 308

Qy 248 -----DSDLWFGYTOQ-----SHWQIFNGKNS-----RPRV----- 274
|| :
Db 309 SQASGDGGPYDEADVWIEFLNENNISWANNSLTN-KNEVSGAFTPEELGKSNATNLDPG 367

Qy 275 --HDYOP--EIFLT-----OPV-----YSDLPW---DGKVRMIGAVHHSNGE 311
|| :
Db 368 PDHVAPEELSLSGEYVRARIKGVNVEPDKTKTKVLWDFNDGKQ---GFGVNSDSNPK 425

Qy 312 SAKLSRSWNRAYLMAGME-----WKNLTVMPIRWGR 342
|| :
Db 426 ELIAVDNENNTLKVSGLDVSDVSDGNFWANARLSANGWK 466

RESULT 15
T20122
hypothetical protein F25B3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T20122; T21324

Search completed: November 30, 2001, 14:18:36
Job time: 258 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:48 ; Search time 36.79 Seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 2363
Sequence: 1 MKVSLSTLTLILPCFAILA.....YNHEATSGVGLMDNMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.5	4.3	1726	2	US-08-609-049A-30
2	100.5	4.3	1726	2	US-09-170-996-30
3	99.5	4.2	1658	4	US-08-609-049A-13
4	99.5	4.2	1658	4	US-09-170-996-13
5	95.5	4.0	682	3	US-08-481-435-6
6	90.5	3.8	857	1	US-07-717-331F-2
7	89.5	3.8	1024	4	US-09-091-117-5
8	88.5	3.7	503	1	US-07-946-497-2
9	88.5	3.7	503	1	US-08-483-322-2
10	88.5	3.7	503	2	US-08-478-882-2
11	87.5	3.7	666	4	US-08-961-083-2
12	86	3.6	781	1	US-08-373-134D-2
13	86	3.6	781	2	US-09-114-637-2
14	86	3.6	985	5	PCT-US96-03916-6
15	86	3.6	985	5	PCT-US96-03916-66
16	85	3.6	816	1	US-07-731-157A-4
17	85	3.6	816	1	US-08-229-444B-2
18	85	3.6	816	2	US-08-541-780-4
19	85	3.6	1627	1	US-07-665-792E-9
20	84.5	3.6	774	3	US-08-902-632-2
21	84.5	3.6	774	3	US-09-073-354-1
22	84.5	3.6	774	3	US-08-656-005A-1
23	84.5	3.6	774	4	US-09-073-259-1
24	84.5	3.6	774	4	US-09-363-095-1
25	84.5	3.6	774	4	US-09-418-027-1
26	84	3.6	522	6	RE34606-6
27	83.5	3.5	657	4	US-09-306-593-2

28	83	3.5	1088	3	US-08-633-768A-1	Sequence 1, Appl
29	82.5	3.5	535	2	US-08-564-972-1	Sequence 1, Appl
30	82.5	3.5	535	4	US-09-171-969-9	Sequence 9, Appl
31	82.5	3.5	797	4	US-09-086-912-2	Sequence 2, Appl
32	82	3.5	320	2	US-08-245-511-4	Sequence 4, Appl
33	82	3.5	320	2	US-08-600-993A-4	Sequence 4, Appl
34	82	3.5	355	1	US-07-946-497-5	Sequence 5, Appl
35	82	3.5	355	1	US-08-483-322-5	Sequence 5, Appl
36	82	3.5	355	2	US-08-478-882-5	Sequence 5, Appl
37	82	3.5	1022	3	US-08-772-270A-2	Sequence 2, Appl
38	82	3.5	1381	4	US-09-540-245A-16	Sequence 16, Appl
39	81.5	3.4	663	1	US-08-441-139-7	Sequence 7, Appl
40	81.5	3.4	693	4	US-08-235-836C-68	Sequence 68, Appl
41	81.5	3.4	844	1	US-07-731-157A-6	Sequence 6, Appl
42	81.5	3.4	844	2	US-08-541-780-6	Sequence 6, Appl
43	81	3.4	334	2	US-08-359-850-4	Sequence 4, Appl
44	81	3.4	537	2	US-08-633-879C-2	Sequence 2, Appl
45	81	3.4	649	4	US-09-618-419-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-609-049A-30
; Sequence 30, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-049A-30

Query Match 4.3%; Score 100.5; DB 2; Length 1726;
Best Local Similarity 18.9%; Pred. No. 0.62;
Matches 86; Conservative 66; Mismatches 145; Indels 159; Gaps 21;
QY 45 QNLEPIDV-----QATOSASTDTANPLDE-----HEPELYTTALENKTMLIN----- 88
DB 563 EDDEAPVLDNKLQYIEKPYKVMTRHPVELDSDHYQVEL-ALQTNQHRVAQVKA 621

Qy 89 -----CSALNODIMLACYDTLVHGETP-----AVIKTKRSIRLDETIMQTIKGPQVYQOE 140
Db 622 VRKICSALED-----GVETPSVTEAVKKLRAVNLPR----- 652
Qy 141 TTDPIFLMGNEKMLTKDKAQLEYAAKQFTPLSLSFOLDNRNTPLMSSRRPHNPMYVLP- 199
Db 653 -----NKSADVT-----SLSGSDTRKNTKGSNLNPNPQVQVSM 686
Qy 200 -----IFMHGKPNR-----SPNTPSHEAKQFTPNFRAPELKFQVSVKVAEDL-- 244
Db 687 HLTRIIDYLLRLHANSRSCSTGCPGRSRNIKEAWTATE-----QLQFTVYAAHGISS 738
Qy 245 -WGTDSDLWFGYTOQSHWQIFNGKN-SRPF---RVHDYQPEIFLTQ-----PV-YSDL 291
Db 739 NWSNYEKYIICLSLH-----NGKDLFKPIQSKKVGTYKKNFFYLIKWDDELIIFFPIQISQL 794
Qy 292 PWDGKVRMIGMCAVHHSNGESAKLSRSWN-----RAYLMAGMEWKNLTVM 336
Db 795 PLESVLHLTLFGVLNQSSGSPDSNKQKQPEALGKVSILTFLDFKRFELTCG-----TKL 848
Qy 337 PRINGRIFKEGSGQPDNDPILDYGYGDVRFYQLENKSNISGTVRYNPRSGKGALQOL 396
Db 849 LYLW-----TSSHTNSIPGAIPKKSVMERIVLQVDFPSPAFDIITYTSPQIDRNIIOQ 901
Qy 397 DYVYPLKGIGSYFOIFQGYGQSLIDYNHEATSRGV 432
Db 902 DKLETLESIDKG-----KLDDIHRDSSFG 927

RESULT 2

US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-170-996-30

Query Match 4.3%; Score 100.5; DB 4; Length 1726;
Best Local Similarity 18.9%; Pred. No. 0.62;
Matches 86; Conservative 66; Mismatches 145; Indels 159; Gaps 21;
Qy 45 QDNELPIDV-----OSATQASTDTANPLDE-----HEPELYTALENKTMNLIN----- 88
Db 563 EDDEAPVDNLKYLQIEKPYKEVMTRHPEVELLDSYHYQVEL-ALQTNQHRADVQVKA 621
Qy 89 -----CSALNODIMLACYDTLVHGETP-----AVIKTKRSIRLDETIMQTIKGPQVYQOE 140
Db 622 VRKICSALED-----GVETPSVTEAVKKLRAVNLPR----- 652
Qy 141 TTDPIFLMGNEKMLTKDKAQLEYAAKQFTPLSLSFOLDNRNTPLMSSRRPHNPMYVLP- 199
Db 653 -----NKSADVT-----SLSGSDTRKNTKGSNLNPNPQVQVSM 686
Qy 200 -----IFMHGKPNR-----SPNTPSHEAKQFTPNFRAPELKFQVSVKVAEDL-- 244
Db 687 HLTRIIDYLLRLHANSRSCSTGCPGRSRNIKEAWTATE-----QLQFTVYAAHGISS 738
Qy 245 -WGTDSDLWFGYTOQSHWQIFNGKN-SRPF---RVHDYQPEIFLTQ-----PV-YSDL 291
Db 739 NWSNYEKYIICLSLH-----NGKDLFKPIQSKKVGTYKKNFFYLIKWDDELIIFFPIQISQL 794
Qy 292 PWDGKVRMIGMCAVHHSNGESAKLSRSWN-----RAYLMAGMEWKNLTVM 336
Db 795 PLESVLHLTLFGVLNQSSGSPDSNKQKQPEALGKVSILTFLDFKRFELTCG-----TKL 848
Qy 337 PRINGRIFKEGSGQPDNDPILDYGYGDVRFYQLENKSNISGTVRYNPRSGKGALQOL 396
Db 849 LYLW-----TSSHTNSIPGAIPKKSVMERIVLQVDFPSPAFDIITYTSPQIDRNIIOQ 901
Qy 397 DYVYPLKGIGSYFOIFQGYGQSLIDYNHEATSRGV 432
Db 902 DKLETLESIDKG-----KLDDIHRDSSFG 927

RESULT 3

US-08-609-049A-13
; Sequence 13, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:

ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.0%; Score 95.5; DB 3; Length 682;
Best Local Similarity 19.9%; Pred. No. 0.45;
Matches 89; Conservative 69; Mismatches 186; Indels 103; Gaps 24;

QY 18 ILAQQAQAVNPVAFVDEVRSENDL-QGDNELPDIQVSATQS-----ASTDTANP 67
DB 10 LVATTSSKIYDNKNQLIADGSESRVNAQANDIPTDLVKAIVSIEDHRRFDRHGDITIRI 69
QY 68 LDEHEPELYTVALENKTNLINCASALNQDMLRACVDTLVHGETPAVTKRSIRLDETIW 127
DB 70 LGAFRLNQSLOGG-----SALTQQLIKLTYSTSTSDOTIS-----RAQEAW 115
QY 128 QTIKGPQVYQET-----TDFIFLMGNEKGMUTK-----DKAKOLE-----YA 166
DB 116 LAIQLEQATKQEIITYINKVYMSNGVGMQTAQNYGKDLNLSLPQLALLAGMPQA 175
QY 167 AKQFTPLS-LSFDDLDRNTPWSSRPHPMYVLPFIMHGKPNRSPNTPSHEAKQFTPNPF 225
DB 176 PNOYDPYSHPEAAQDRRLNLV--SEMKNOGYI-----SAEQYEKAVNTPTDGLQSLKSA 229
QY 226 RAPE-----LKFQVS-VKVKAAEDLWGTDSLWFGYTO--QSH-WQIFNGKNSRFRVHD 276
DB 230 NYPAYMDNLYKEVINQVEETGYNLTTGMDVYTNVQEAQKHLMDIYNTDEYVAYPDE 289
QY 277 YQPEIFLQPVYSDLPWBDGKVRMIGMGAHVHRSN-----GESAKLSRSWNRAYLMAGMEW 330
DB 290 LQ-----VASTIVDVS-NGKV-IAQLGARHQSNNVSYFGINQAVETNRD-----GSTM 335
QY 331 KNLT-VMPRIKGRIPKESGGSDPDNDILDYQ-----YGDVRFYOLENKS 377
DB 336 KPITADYAPALEYGVY-ESTATIVHDEP--YNYPGTNTPVNMDRGYFGNITLQYALQOSR 392
QY 378 NISGTVRYN-----PRSGKALQLDY 398
DB 393 NYPAVETLNKVLGNRAKFTLNLGLIDY 419

RESULT 6

US-07-717-331F-2
; Sequence 2, Application US/07717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07717.331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-717-331F-2

Query Match 3.8%; Score 90.5; DB 1; Length 857;
Best Local Similarity 21.4%; Pred. No. 2.2;
Matches 76; Conservative 50; Mismatches 110; Indels 119; Gaps 23;

QY 105 LVHGETPAVTKRSIRLDETIWQIKGPOVYQETTDPIFLMG-----NEK---GMLTK 157
DB 24 LIHPALSYINT-----LSSTESLTISNKTLY-----SPGSIFEVGFRTNSRWYLGMMYK 75
QY 158 K-DAKOLEYAAKQFTPLSLSFED-----LDRNNTPLWSSRPHNPMYVLPFIMHGK 205
DB 76 KVSRTYVAVANRDNPLSNAIGTLAISGNLVLLDHSHKPNVWNTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAKQFT-----PNEFRAPELKFQVSVKVKAAEDLW 245
DB 126 -ERSPPVAELLANGNFVNRDSSNDASEVLMQSFYDPTDILLPEMKLYNLT-----177
QY 246 GTDSDLWFGYQOOSHWQIFNGKNSRPFVRHDYQ-----PEIFLQO---PVYSDLPWQK 296
DB 178 GLNRF-----TSWRSSDDPSGNGF---SYKLEQSLPEFLSRENFPMHRSGPWNG- 226
QY 297 VRMIGMGAVH-----HSNGESAKLSRSWNRAY-----LMAGMEWKNLTVM---RI 339
DB 227 IRFSGIPEDQKLSYVYFNFIENNEEVAYFTNTNSFVSRLLTISEGYFQRLTWPSIRI 286
QY 340 WGRIFKEGSGQDD-----NPDILDYGYGDVRFYQLENKSNISGTVR-YNPRS 389
DB 287 WNRFWSSPDVQCDYFIMCGP-----YAYCDV-----NTSPVCNCIOGFNPRN 329

RESULT 7

US-09-091-117-5


```

; Sequence 5, Application US/09091117
; Patent No. 6171589
; GENERAL INFORMATION:
; APPLICANT: The University of Melbourne
; TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
; TITLE OF INVENTION: Vaccines
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: United States of America
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,117
; FILING DATE: 12 JUNE 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU96/00803
; FILING DATE: 13-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN7127
; FILING DATE: 13-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WINNER, Ellen P.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: +1 303 499 8080
; TELEFAX: +1 303 499 8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1024 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Mycoplasma genitalium
;
US-09-091-117-5

```

```

Query Match      3.8%; Score 89.5; DB 4; Length 1024;
Best Local Similarity 20.2%; Pred. No. 3.7;
Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

QY 1 MKVS-LSTLTSLPCFALLATQQAQVNPVAFVDEKSENLDGDNELPI-DVQSATQ 58
Db 1 MKLSTTTTCLTSISGAFGTTAI----ALPTTVALLNKHHQQNTEKQON--PIKDIFRGLN 54

QY 59 SASDTANPLDHEPELYTTALENKTMLNCSALNQDINRLACYDTLVHGETPAVTKR 118
Db 55 NVQVPNTIPL--HQTVEVT--NNKAIVDYKDAPOKFFL-----AKSALNNKL 98

QY 119 SIRLDETIMQTIKGRPOVVYQETTDPIFLMGNEKGLTKKDAKQLEYAAKO-FTPLSLSF 177
Db 99 QVEFDKFLRT-----GVINALNADLKEWIDQTLFIPNQSF 135

QY 178 DLDRNNTPLWSSRPINPMVLP-IFMHGKPNRP-NTPSHEAKQFTPNFRAPELKFQVS 235
Db 136 DLSANKNLNLTLSQSEVSLDLEFIITFNSDKNQPLKLPFGDSVVVWANE---SYTYSVK 191

QY 236 VKVKAEDLWGTSDSLWFGYQTOQSHWQIFNGKNSRPFVRHYDQPEI-FLTPQVYSDL-PW 293
Db 192 ATLQKLKLVITYSRADHSVGSIAIPTVSLNGTKQNDQSFNPFKSNINFAKVNYYNALNPF 251

QY 294 DGKVRMIGMGAVHVS--NGESAK-----LSRSWNRAYLMAGMENKLVMPRIWGRIEK 345
Db 252 EAQQVLVGQKGLFNQKNVADDDKNDINNHIETQFNVAKITA-----TLGKAKFK 300

QY 346 ---EGSGSQP-----DDNPDILDY--GYGDV-----RFLYQL--- 373

```

```

Db 301 QFGEHKNQGPLSLKLVSLGNNFEKOLFVNYVRPGLDFVSLQSSQSSNKKTVYQLLF 360
QY 374 ENKSNI 379
Db 361 ENKTTI 366

RESULT 8
US-07-946-497-2
; Sequence 2, Application US/07946497
; Patent No. 5506119
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHER, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-07-946-497-2

Query Match      3.7%; Score 88.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.5;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

QY 40 ENDLGDNELPIDVQSATQSASDTANPLDHEPELYTTALENKTMLNCSALNQDINRL 99
Db 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYLHTDLPT---SQPTGDRD--- 212

QY 100 ACVDTLVHGETPAVTKRSI-----RLDETWTQTIKGRPOVVYQETTDPIFLMGNEK 153
Db 213 ---DAFFIGSTLATIATTPWNSAHTKQNTQWNIHNSPEVLLQTTT----- 258

QY 154 MLTKDAKQLEYAAKQFTPLSLGSFIDLDRNNTP-----WSSRPINPMVLP-IFMHGKPNRS 209
Db 259 -----RMT-----DIDRNSTSARGENWNTQEPQPF----- 283

QY 210 PNTPSHEAKQFTPNFRAPELKFQVSVKVAEEDLW-----GTDSDLWFGYTTQS 259
Db 284 -NNHEYODEETPH-----ATSTTWADPNSTTEAATQKEKWF-----EN 322

```

Qy 260 HWQIFNGKN-----SRPRVHDYQEIFLTPVYSDLPW-----DGKVRMIGMG 303
 Db 323 EQW---GKNPTPSDSHVTEGTTASAHNNHPSQRMWTTQSQEDVSWTDFDFISHPMGQG 379
 Qy 304 AVHNSGESA 313
 Db 380 HQTESKGHS 389

RESULT 9
 US-08-483-322-2
 ; Sequence 2, Application US/08483322
 ; Patent No. 5760178
 ; GENERAL INFORMATION:
 ; APPLICANT: HERRLICH, Peter
 ; APPLICANT: PONTA, Helmut
 ; APPLICANT: GUENTHER, Ursula
 ; APPLICANT: MATZKU, Siegfried
 ; APPLICANT: WENZL, Achim
 ; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
 ; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
 ; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/483.322
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/946,497
 ; FILING DATE: 09-NOV-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16915/145
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 503 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-483-322-2

Query Match 3.7%; Score 88.5; DB 1; Length 503;
 Best Local Similarity 18.4%; Pred. No. 1.5;
 Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

Qy 40 ENDLGQDNEPLDVQSATQASDTANPLDHEPELYTTALENKTMLNCSALNODIMRL 99
 Db 169 QEDIDASNIIDEDVSS-----GSTIEKSTPEGYLLHTDLP-----SQTGDRD----- 212

Qy 100 ACYDTLVHGETPAVIKTKRSI-----RLDETIWQTIKGPQVYQETTDPIFLMGNEKG 153
 Db 213 ---DAFFIGSLATATPPWISAHTKQKQERTQWNPPIHNSPEVLLQTT----- 258

Qy 154 MLTKKDAKQLEYAAKQFTPLSLSDLDNRNTP-----WSSRPHNPVYLPIFMHGKPNRS 209

Db 259 -----RMT-----DIDRNSTSAHGENWNTQEPQPP----- 283
 Qy 210 PNTPSHEAKQFTNPEFRAPELKQVSVKKAEDLW-----GTDSDLWFCYTOQS 259
 Db 284 -NNHEYQDEETPH-----ATSTTWADPNSTTEAAATQKEKWF-----EN 322
 Qy 260 HWQIFNGKN-----SRPRVHDYQEIFLTPVYSDLPW-----DGKVRMIGMG 303
 Db 323 EQW---GKNPTPSDSHVTEGTTASAHNNHPSQRMWTTQSQEDVSWTDFDFISHPMGQG 379
 Qy 304 AVHNSGESA 313
 Db 380 HQTESKGHS 389

RESULT 10
 US-08-478-882-2
 ; Sequence 2, Application US/08478882
 ; Patent No. 5885575
 ; GENERAL INFORMATION:
 ; APPLICANT: HERRLICH, Peter
 ; APPLICANT: PONTA, Helmut
 ; APPLICANT: GUENTHER, Ursula
 ; APPLICANT: MATZKU, Siegfried
 ; APPLICANT: WENZL, Achim
 ; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
 ; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
 ; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/478.882
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/946,497
 ; FILING DATE: 19921109
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16915/145
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 503 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-478-882-2

Query Match 3.7%; Score 88.5; DB 2; Length 503;
 Best Local Similarity 18.4%; Pred. No. 1.5;
 Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

Qy 40 ENDLGQDNEPLDVQSATQASDTANPLDHEPELYTTALENKTMLNCSALNODIMRL 99
 Db 169 QEDIDASNIIDEDVSS-----GSTIEKSTPEGYLLHTDLP-----SQTGDRD----- 212

Qy 100 ACYDTLVHGETPAVIKTKRSI-----RLDETIWQTIKGPQVYQETTDPIFLMGNEKG 153

Db 213 ---DAFFIGSTLATTTPWSAHTKQONERTQWNP IHSNPEVLLQTTT----- 258
QY 154 MLTKDKAKQLEAAKQFTPLSLSFDRNNTPL---WSSRPHNPMYVLP IFMHGKPNRS 209
Db 259 -----RMT-----DIDRNSTSAHGENWTQEPQPP----- 283
QY 210 PNTPSHAKQFTNEFRAPELKFQVSVKYKAAEDLW-----GTDSDLWFGYTOQS 259
Db 284 -NNHEYQDEETPH-----ATSTTWADPNSTTEEAATQKEKF-----EN 322
QY 260 HQIFNGKN-----SRPRVHDYQPEIFLTQPVYSDLPW-----DGKVRMIGHG 303
Db 323 EWO---GKNPPTSEDSHVTEGTASAHNHPQRSQRTTQSDQEDVSWTDFDPISHPMGOG 379
QY 304 AVHHSNGESA 313
Db 380 HQTESKGHSS 389

RESULT 11
US-08-961-083-2
; Sequence 2, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-2

Query Match 3.7%; Score 87.5; DB 4; Length 666;
Best Local Similarity 20.4%; Pred. No. 3;
Matches 87; Conservative 60; Mismatches 176; Indels 103; Gaps 24;
QY 42 DLG-----QNELPIDVQSATQS-----ASTDTANPLDEHEPELYTTALENKTM 85
Db 12 DLGSRVNAQANDITDLVKAIVSIEDHRFDHGRIDIRILGAFRLNLSLOGG--- 69
QY 86 LINCALSALQDMLRLACYDTLVHGETPAVTKRSIRLDETINQTIKGPQVYQET---- 141

Db 70 ---STLTQOLIKLYFTSTSDOTIS-----RKAQEAWLAIQLEQKATKOEILTY 117
QY 142 TDFIFLMGNEKGLTK-----KDAKQLE-----YAAKQFTPLS-LSFOLDNRN 183
Db 118 INKVMNSNGYMQTAQNYGKDLNNLSLPQJALLAGHPQAPNQDYPYSHPEAAQDRN 177
QY 184 TPLWSSRPHNPMYVLP IFMHGKPNRSPTPSHEAKQFTPNFRAPE-----LKFEQVS-VK 237
Db 178 LVL--SEMKNQGYI---SAEQYEKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQVE 231
QY 238 VKAAEDLWGTDSDLWFGYTO---QSH-WQIFNGKNSRPRVHDYQPEIFLTQPVYSDLPW 294
Db 232 EETGYNLLTGMVYTVNDQEAQKHLMDIYNTDEYVAYPDDELQ-----VASTIVDVS-N 285
QY 295 GKVRMIGMGAVHHSN-----GESAKLSRSWNRAYLMAGMEKNLT-VMPRI-----W 340
Db 286 GKV-IAQLGARHQSNNVSTGNGAVETNRDW-----GSTMKPTIDYAPALEGYVDST 337
QY 341 GRIFKEGSGQDDNDPILDY-YGY-GDVRFLYQLENKSNISGTVRYN-----PRSGKG 392
Db 338 ATIVHDEPYNPGTNPVYVWMDRGYFGNITLQVALQOSRNVPAVETLNKVLNRAKTEFN 397
QY 393 ALQLDY 398
Db 398 GLGIDY 403

RESULT 12
US-08-373-134D-2
; Sequence 2, Application US/08373134D
; Patent No. 5780296
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric
; APPLICANT: Holloman, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
; TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,134D
; FILING DATE: January 17, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Friebe, Thomas E.
; REGISTRATION NUMBER: 29,258
; REFERENCE/DOCKET NUMBER: 7991-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 781 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-373-134D-2

Query Match 3.6%; Score 86; DB 1; Length 781;
Best Local Similarity 20.7%; Pred. No. 5.5;

QY 107 HGETPAVITKRSIRLDETIWOTIKGKPOVVYOE-TTDP1---FLMGNEKGM1TKKDAKQ 162
Db 614 FTQAPSTVPRKAT-----QTPSTEPEVLTAQSPSTEPVPTRTILGAPEITQTPSA-- 662
QY 163 LEVAAKQFTPLSLFDRNNTPLWSSRP-----HN-----PMVLP-----IFMHGK 205
Db 663 ---APEVYTRSSSTMPETAQSTPLASQNTSSGTHNTEPRTPYVQTTPHQKLYTENK 719
QY 206 PNRSPNTPS--HE---AKQFTP-NEFRAPELKFOVSVKVKYK-----AEDLWGTD 248
Db 720 TLSFPTVWSEFHEMSTAESQTPLLDKIVKVEKFSNDGEVATCVSVKSPYRVETNNKVD 779
QY 249 -----SDLWFGYTQOSHWOIFNGKNS-----RPF----- 272
Db 780 LVDVMEISGNSPAGVFNSEKRW---OKQLYYRVTDGRTSVQMLCLSLCTSHSPEYCLFD 836
QY 273 -----RVHDYQPEIFLT---QPVYSDLPWDGKV-----RMIGMGAVHHNS 309
Db 837 TSLIAREKDIAPELYFTSDPQAYCTITLPSGVVPRFENSLNNVSLPEYLTATTVVSHTA 896
QY 310 GE-----SAKLSRSW 319
Db 897 GQSTVWKSSARAGEAW 912

RESULT 15
PCT-US96-03916-66
; Sequence 66, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-03916-66

Query Match 3.6%; Score 86; DB 5; Length 985;
Best Local Similarity 19.4%; Pred. No. 8;
Matches 73; Conservative 47; Mismatches 126; Indels 130; Gaps 20;

QY 48 ELPIDVQSATQTSASTDTANPLDEHEPELYT-TALENKTMLINCNSALNQDIMRLACVDYTLV 106
Db 563 QTPSTVPKTTLSST-----EPAIFTRTQSGTEAFTQTSSAEPOTMRTQSTETHF 613
QY 107 HGETPAVITKRSIRLDETIWOTIKGKPOVVYOE-TTDP1---FLMGNEKGM1TKKDAKQ 162
Db 614 FTQAPSTVPRKAT-----QTPSTEPEVLTAQSPSTEPVPTRTILGAPEITQTPSA-- 662
QY 163 LEVAAKQFTPLSLFDRNNTPLWSSRP-----HN-----PMVLP-----IFMHGK 205
Db 663 ---APEVYTRSSSTMPETAQSTPLASQNTSSGTHNTEPRTPYVQTTPHQKLYTENK 719
QY 206 PNRSPNTPS--HE---AKQFTP-NEFRAPELKFOVSVKVKYK-----AEDLWGTD 248
Db 720 TLSFPTVWSEFHEMSTAESQTPLLDKIVKVEKFSNDGEVATCVSVKSPYRVETNNKVD 779
QY 249 -----SDLWFGYTQOSHWOIFNGKNS-----RPF----- 272
Db 780 LVDVMEISGNSPAGVFNSEKRW---OKQLYYRVTDGRTSVQMLCLSLCTSHSPEYCLFD 836
QY 273 -----RVHDYQPEIFLT---QPVYSDLPWDGKV-----RMIGMGAVHHNS 309
Db 837 TSLIAREKDIAPELYFTSDPQAYCTITLPSGVVPRFENSLNNVSLPEYLTATTVVSHTA 896
QY 310 GE-----SAKLSRSW 319
Db 897 GQSTVWKSSARAGEAW 912

Search completed: November 30, 2001, 14:17:51
Job time: 228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: November 30, 2001, 14:17:00 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec
Title: US-09-787-083-8
Perfect score: 2363
Sequence: 1 MKVSLSTLTLILPCFAILA.....YNHEATSGVGLMLNDWML 442
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	2363	100.0	442	21	AA1985271
2	2354	99.6	442	21	AA1985268
3	2348	99.4	442	21	AA1985270
4	2339	99.0	442	21	AA1985269
5	822	34.8	370	21	AA1975156
6	815	34.5	370	21	AA1975157
7	815	34.5	370	21	AA1975158
8	815	34.5	374	21	AA1970629
9	812	34.4	375	21	AA1970628
10	246.5	10.4	355	19	AA198871
11	239.5	10.1	356	19	AA198960

12	156.5	6.6	253	18	AAW20760	H. pylori outer me
13	117	5.0	1686	19	AAW70991	Human class II P13
14	108.5	4.6	800	8	AAW70420	Sequence encoded b
15	108.5	4.6	822	13	AAW26021	Alkaline cellulase
16	107.5	4.5	824	21	AAW23180	Bacillus sp. KSM-S
17	107	4.5	157	18	AAW20538	H. pylori outer me
18	106	4.5	798	21	AAW40925	Human ORF689
19	106	4.5	798	21	AAW90225	Human GGC-1 protei
20	106	4.5	798	22	AAW83944	Peroxisome prolif
21	102.5	4.3	537	22	AAW91000	C glutamicum prote
22	100.5	4.3	1726	18	AAW38756	Phosphatidyl inos
23	98.5	4.2	467	20	AAW13378	Amino acid sequenc
24	98.5	4.2	467	21	AAW94860	Human protein clon
25	98.5	4.2	467	22	AAW88401	Human membrane or
26	98.5	4.2	467	22	AAW80246	Human PRO260 prote
27	95.5	4.0	502	22	AAW90031	C glutamicum prote
28	95.5	4.0	516	22	AAW78876	C. glutamicum SRT
29	95.5	4.0	682	17	AAW04359	S. pneumoniae penl
30	95	4.0	522	22	AAW95616	Human protein sequ
31	95	4.0	888	22	AAW70751	S cerevisiae apopt
32	94	4.0	372	16	AAW69607	Neisseria meningit
33	94	4.0	761	20	AAW99084	Non-B, non-C, non-
34	94	4.0	1227	22	AAW81501	S. epidermidis ope
35	94	4.0	1717	22	AAW20498	Ostrinia nubilalis
36	93.5	4.0	467	19	AAW75057	Human colon cancer
37	93	3.9	600	22	AAW75092	S receptor kinase
38	92.5	3.9	857	13	AAW29814	Cytadhesin P1. My
39	91	3.9	1627	16	AAW67538	Human protein sequ
40	90	3.8	522	22	AAW95513	Modified dipteria
41	90	3.8	560	11	AAW06815	T. matsutake pyran
42	90	3.8	564	21	AAW10457	Trichoderma deri
43	90	3.8	564	21	AAW81952	Human protein sequ
44	90	3.8	790	22	AAW93045	Mouse dephosphoryl
45	90	3.8	841	20	AAW24318	

ALIGNMENTS

RESULT 1	
AAW85271	AAW85271 standard; Protein; 442 AA.
XX	AC
XX	AC
XX	29-JUN-2000 (first entry)
XX	XX
XX	BASB034 amino acid sequence #4.
XX	XX
XX	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	XX
OS	Moraxella catarrhalis.
XX	XX
PN	WO200015802-A1.
XX	XX
PD	23-MAR-2000.
XX	XX
PF	14-SEP-1999; 99WO-EF06781.
XX	XX
XX	14-SEP-1998; 98GB-0020002.
XX	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PI	Ruelle J;
DR	WPI; 2000-271440/23.
DR	N-PSDB; AAA10703.
XX	XX
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
PT	catarrhalis used to prepare vaccines against bacterial infections

XX PS Claim 3; Page 69; 106pp: English.

XX CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

XX CC strain Mc2969. The invention relates to BASB034 polypeptides from

XX CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034

XX CC polynucleotides and polypeptides may be employed as research reagents and

XX CC material for the discovery of treatments and diagnostics for diseases,

XX CC particularly human diseases. They are particularly used to diagnose and

XX CC treat M. catarrhalis infections. They can be used for diagnosis of

XX CC disease, staging of disease, or determining response of an infectious

XX CC organism to drugs. The polynucleotides may be used as a source for

XX CC hybridization probes, and for screening of genetic mutations, serotype,

XX CC organism or strain identification, identification of mutations in BASB034

XX CC sequences, and as components of arrays which are useful for diagnostic

XX CC and prognostic purposes. The polypeptides can be used to produce

XX CC antibodies. The polypeptides can also be used in vaccine formulations,

XX CC and to identify agonists and antagonists. The polypeptides, antibodies,

XX CC agonists and antagonists (which are bacteriostatic) are used for the

XX CC treatment and prevention of diseases such as otitis media in infants and

XX CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

XX CC invasive diseases, and chronic otitis media with hearing loss. The

XX CC polypeptides, agonists and antagonists are also used for screening of

XX CC antibacterial drugs. The BASB034 products of the invention can be used

XX CC screen for new antibacterial compounds that may target resistant

XX CC bacteria.

XX SQ Sequence 442 AA;

Query Match 100.0%; Score 2363; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 9.3e-219;
 Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVSLSTLTSLPCFALLATQQAQAVNPVAFVDEVSENDLGDNELPIDVQSATQSA 60
 Db 1 mkvslstlslpcfallatqqaqavnpvafdevrsendlgdnelpidvgsatqsa 60

Qy 61 STDANPLDEHEPELYTTALENKTMLINCSALNQDIMRLACYDTLVHGETPAVTKTRSI 120
 Db 61 stdtanpldehepelyttalenktmlincsalnqdimrlacydtlvhgetpavtktrsi 120

Qy 121 RLDETIWQITKGPQVYQETTDPIFLMGNEKGMITKKDAKOLEYAAKQFTPLSLSFOLD 180
 Db 121 rldetiwtikgpkqvvyqettdpiflmgnekgmtkkdakoleyaaakqftplsfsfld 180

Qy 181 RNTPLWSRRPHNPMYVLPIMPHGKPNRSPNTPSHEAKQFTPNRPRAPLQVSKVKA 240
 Db 181 rntplwsrrphnpmvlpimphgkpnrsntpsheakqftpnrprapelfqvskvka 240

Qy 241 AEDLWGTDSLDLWFGYQTSQHWQIFNGKNSRPRFRVHDYQPEIFLTPQVYSDLPWDGKVRMI 300
 Db 241 aedlwgtdsldlwfgytqshwqifngknsprrfrvhdypelfltpqvysdlpwdgkvrm 300

Qy 301 GMGAVHHSNGESAKLSRSWNRAYLMAGMEWKNLTVMYPRIGRIFKESGSGQPDNDPILD 360
 Db 301 gmgavhhsngesaklsrswnraylmagmewknltympriwgrifksgsgqpdndpild 360

Qy 361 YYGVDVRFLEOLENKNISGTVRNPRSGKALQLDVYVPLGKISGISYFQIFQYGGSL 420
 Db 361 yygvdvrflyqlenknisgtvrynrpsrgkalqldvvyplgkigsgisfyfifqyggsl 420

Qy 421 IDYNHEATSFVGGLMDNMGL 442
 Db 421 idynheatsfvgglmdnmgl 442

RESULT 2
 AAY85268
 ID AAY85268 standard; Protein: 442 AA.
 XX
 AC AAY85268;
 XX

DT 29-JUN-2000 (first entry)
 XX BASB034 amino acid sequence #1.
 DE
 XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX
 XX WO200015802-A1.
 PN
 XX
 XX 23-MAR-2000.
 PD
 XX
 PF 14-SEP-1999; 99WO-EP06781.
 XX
 PR 14-SEP-1998; 98GB-0020002.
 XX
 PA (SMIK) SMITHLINE BEECHAM BIOLOGICALS.
 XX
 PI Ruelle J;
 XX
 DR WPI; 2000-271440/23.
 DR N-PSDB; AAA10700.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections -
 XX
 PS Claim 3; Fig 2; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

SQ Sequence 442 AA;

Query Match 99.6%; Score 2354; DB 21; Length 442;
 Best Local Similarity 99.5%; Pred. No. 6.8e-218;
 Matches 440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSLSTLTSLPCFALLATQQAQAVNPVAFVDEVSENDLGDNELPIDVQSATQSA 60
 Db 1 mkvslstlslpcfallatqqaqavnpvafdevrsendlgdnelpidvgsatqsa 60

Qy 61 STDANPLDEHEPELYTTALENKTMLINCSALNQDIMRLACYDTLVHGETPAVTKTRSI 120
 Db 61 stdtanpldehepelyttalenktmlincsalnqdimrlacydtlvhgetpavtktrsi 120

Qy 121 RLDETIWQITKGPQVYQETTDPIFLMGNEKGMITKKDAKOLEYAAKQFTPLSLSFOLD 180
 Db 121 rldetiwtikgpkqvvyqettdpiflmgnekgmtkkdakoleyaaakqftplsfsfld 180

181 RNTPLWSSRRPNMYVLP IFMHGKPNRSPNTPSHEAKQFTNEFRAPELKFOVSVKVA 240
 181 rntplwssrrpnhmyvlp ifmhgkpnrsntpspsheakqftnefrapelkfqvsvkva 240
 241 AEDLWGTSDLMFQGTQSHWQIFNGKNSRPRVHDYQPEIFLTQPVYSDLPWDGKVRMI 300
 241 aedlwgtstdlmfqtqshwqifngknsrprvrhdyqpeifltqpvysdlpwdgkvrm 300
 301 GMGAVHHSNGESAKLSRSNNRAYLMAGMEWKNLTVMPTWGRIFKEGSGSQPDNDPILD 360
 301 gmgavhhsngesaklsrsnnraylmagmewknlvtmptwgrifkegsgsqpddndpild 360
 361 YGYGDBVRFYQLENKSNISGTVRYNPRSGKALQLDVYVPLGKISGYFOIFQGYGQSL 420
 361 ygygdbvrflyqlenknsisgtvrynprrsgkalqldvypplgkigsgyfoifqgygqsl 420
 421 IDYNHEATSFVGGLMLNDWMGL 442
 421 idynheatsfgvglmndwmgl 442

RESULT 3
 ID AAY85270 standard; Protein; 442 AA.
 AC AAY85270;
 DT 29-JUN-2000 (first entry)
 XX BASH034 amino acid sequence #3.
 KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX
 PN WO200015802-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 14-SEP-1999; 99WO-EP06781.
 XX
 PR 14-SEP-1998; 98GB-0020002.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Ruelle J;
 XX
 DR WPI: 2000-271440/23.
 DR N-PSDB; AAA10702.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections
 XX
 PS Claim 3; Page 68; 106pp; English.
 XX
 CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 strain Mc2913. The invention relates to BASB034 polypeptides from
 M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX Sequence 442 AA;
 SQ
 Query Match 99.4%; Score 2348; DB 21; Length 442;
 Best Local Similarity 99.3%; Pred. No. 2.6e-217;
 Matches 439; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTLSILPCFAILAIQAOQAVPNPFAVDFVRSNDLGDNELPIDVQSQTOSA 60
 DB 1 mkvslstltlsilscfaillaiqqakavpnpvafvdevrsndlggnelpdvqsatqsa 60
 QY 61 STDTANPLDEHEPELYTTALENKTMLINGSALNQDLMRLACYDTLVHGETPVIKTKRSI 120
 DB 61 stdtanpldehepeelyttalenktmlinsalnqdimrlacydtlvhgetpaviktksr 120
 QY 121 RLDETITWQTIKGPQVYQETDPIFLMGNEKGMLTCKDAKOLEYAAKQFTPLSLSFDDL 180
 DB 121 rldetiwtikgkpqvvyqettdpiflmgnekgmickdakqleyaakqftplsld 180
 QY 181 RNTPLWSSRRPNMYVLP IFMHGKPNRSPNTPSHEAKQFTNEFRAPELKFOVSVKVA 240
 DB 181 rntplwssrrpnhmyvlp ifmhgkpnrsntpspsheakqftnefrapelkfqvsvkva 240
 QY 241 AEDLWGTSDLMFQGTQSHWQIFNGKNSRPRVHDYQPEIFLTQPVYSDLPWDGKVRMI 300
 DB 241 aedlwgtstdlmfqtqshwqifngknsrprvrhdyqpeifltqpvysdlpwdgkvrm 300
 QY 301 GMGAVHHSNGESAKLSRSNNRAYLMAGMEWKNLTVMPTWGRIFKEGSGSQPDNDPILD 360
 DB 301 gmgavhhsngesaklsrsnnraylmagmewknlvtmptwgrifkegsgsqpddndpild 360
 QY 361 YGYGDBVRFYQLENKSNISGTVRYNPRSGKALQLDVYVPLGKISGYFOIFQGYGQSL 420
 DB 361 ygygdbvrflyqlenknsisgtvrynprrsgkalqldvypplgkigsgyfoifqgygqsl 420
 QY 421 IDYNHEATSFVGGLMLNDWMGL 442
 DB 421 idynheatsfgvglmndwmgl 442

RESULT 4
 AAY85269
 ID AAY85269 standard; Protein; 442 AA.
 XX
 AC AAY85269;
 DT 29-JUN-2000 (first entry)
 XX
 DE BASB034 amino acid sequence.#2.
 XX
 KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX
 PN WO200015802-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 14-SEP-1999; 99WO-EP06781.
 XX

PR 14-SEP-1998; 98GB-0020002.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Ruelle J;
 XX WPI; 2000-271440/23.
 DR N-PSDB; AAA10701.
 XX
 XX Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections -
 XX
 PS Claim 3; Page 67; 106pp; English.
 XX
 CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX
 SQ Sequence 442 AA;

Query Match 99.08; Score 2339; DB 21; Length 442;
 Best Local Similarity 99.18; Pred. No. 1.9e-216;
 Matches 438; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSLSTLTSLPCFALLAQQAQVNPVAFVDEVSENDLGQDNELPDIQVQSATQSA 60
 Db 1 mkvslstltlslpcfallaiaqqagvnpvafvdevskndlgqdnelllgvqsatqsa 60

Qy 61 STDTPANPLDEHEPELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVINKTRSI 120
 Db 61 stdtanpldehepeylttalenktmlincsalnqdimrlacydtlvhgetpavinktrsi 120

Qy 121 RLDETINQTIKGGQVYQETTDIFLGNKEKGMTTKDAKQLEYAAKQFTPLSLFSLD 180
 Db 121 rldetinqtikgqvvyqettdpiflmgnekgmttkdakqleyaakqftplsld 180

Qy 181 RNNTPLSSRRPHNPMYVLPFIMHGKPNRSPNTPSHEAKOFTNPEFRAPELFQVSVKVA 240
 Db 181 rntplssrrphnmyvlpfimhgkpnrsptpshearkoftnpefrapelfqsvkvka 240

Qy 241 AEDLWGTSDSLWFGYTOQSHQWQIFNGKNSRFRVHDYOPEIFLTPVYSDLPWQKVRMI 300
 Db 241 aedlwgtdslwfgytoqshqwqifngknsrfrvhdypelfltpvysdlpwqkvrm 300

Qy 301 GMGAVHNSGESAKLSRSWNPAYLWAGNWKNTVMPRIWGRIFKEGSGSQPDNDPDL 360
 Db 301 gmgavhnsgegaklsrswnpaylmagnewkntvmprirwgrifkegsgsqpdndpdl 360

Qy 361 YYGVDVRFYQLENKSNISGTQVYRNPRSGKALQDLYVYPLGKISGYFOIFQYQOSL 420
 Db 361 yygvdvrfyqlenksnisgtvynprsgkalqldlyvypkgisgyfoifqyqosl 420

Qy 421 IDYNHEATSFVGLMLNDWMGL 442
 Db 421 idynheatsfvgvlnldwmgl 442

RESULT 5
 AAY75156
 ID AAY75156 standard; Protein; 370 AA.
 XX
 AC AAY75156;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ53918.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 903; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 370 AA;

Query Match 34.8%; Score 822; DB 21; Length 370;
 Best Local Similarity 44.7%; Pred. No. 1.6e-70;
 Matches 163; Conservative 68; Mismatches 110; Indels 24; Gaps 8;

Qy 87 INCSALNODIMRLACYDTLVHGETPAVI-----KYKRSIRLDETIWQTI-KGKQVQVQE 140

Db	193	sdwqinyngqrksapfntdykpeifltqpvkadtlfgrlrmलगsfvqhngngsrpear	252
Qy	318	SNWRAYLMAGMEWKNLTVMPRWGRIFKEGSGSQDDNDPIDLYYGYGDVRFLLQLENKS	377
		: : : : : :	
Db	253	snwriyamagewgkltvprwvrafdq-sgdk-ndnpdiadymgydvklqyrlndrq	310
Qy	378	NISGTVRYNPRSKGALQLDYVYVPLKGGLSGYFQIFQGYGOSLDIYNHEATSFVGGLMLN	437
		: : : : : :	
Db	311	nvysvlrpnkptgygaleaaytfpkglkvrvrgfhgygesldynhkhqngigiglmfn	370
Qy	438	DWMGL 442	
Db	371	dlldgi 375	
RESULT	10		
AAW98871			
ID	AAW98871	standard; Protein; 355 AA.	
XX	AC	AAW98871;	
XX	XX		
DT	31-MAR-1999	(first entry)	
XX	XX		
DE	H. pylori	GHPO 1723 protein.	
XX	XX		
KW	GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;		
KW	peptic ulcer disease.		
XX	XX		
OS	Helicobacter pylori.		
XX	XX		
PN	WO9843478-A1.		
XX	XX		
PD	08-OCT-1998.		
XX	XX		
PF	01-APR-1998; 98WO-US06371.		
XX	XX		
PR	29-JUL-1997; 97US-0902615.		
01-APR-1997;	97US-0833457.		
PR	24-JUN-1997; 97US-0881227.		
XX	XX		
PA	(HUMA-) HUMAN GENOME SCI INC.		
PA	(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.		
XX	XX		
PI	Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;		
XX	XX		
DR	WPI; 1998-542293/46.		
DR	N-PSDB; AAX14590.		
XX	XX		
PT	New isolated Helicobacter polynucleotides - used to develop products		
PT	for the diagnosis, prevention and treatment of Helicobacter		
PT	infections and gastrointestinal diseases		
XX	XX		
PS	Claim 8; Page 1976-1977; 2054pp; English.		
XX	XX		
CC	This sequence represents a Helicobacter pylori GHPO protein of the		
CC	invention. The polypeptides can be used for preventing or treating		
CC	Helicobacter infections, and gastroduodenal diseases associated with		
CC	these infections, including acute, chronic, and atrophic gastritis, and		
CC	peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be		
CC	used for the production of antibodies. The products can also be used for		
CC	detection and diagnosis.		
XX	XX		
SQ	Sequence 355 AA;		
Query Match	10.4%; Score 246.5; DB 19; Length 355;		
Best Local Similarity	25.1%; Pred. No. 3.8e-15;		
Matches	82; Conservative 41; Mismatches 117; Indels 87; Gaps 11		
Qy	157	KDAKQLEYAAKQFTPLSLSFOLDLRNNTPLWSSRPHNPMYVLPFPMHCKPNRSPTPSHE	216
		: : : : : : : : : : : : : : :	
Db	69	kkylnmndyigtvyflpfyhsf-----tpifqwyhpnlnp-----	102

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:28 ; Search time 77.9 Seconds
(without alignments)
829,941 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 2363
Sequence: 1 MKVSLSTLTLILPCFAILA.....YNHEATSFVGVLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2 Q9K0U7	Q9K0U7 neisseria m
2	815	34.5	409	2 Q3JT21	RT Q3JT21 neisseria m
3	372	15.7	306	2 Q9CL22	Q9CL22 pasteurella
4	366	15.5	329	2 Q9PMU8	Q9PMU8 campylobact
5	360.5	15.3	292	2 Q9SLD7	Q9SLD7 yersinia ps
6	360	15.2	289	2 Q9L6N9	Q9L6N9 salmonella
7	356.5	15.1	292	2 Q9Z4N8	Q9Z4N8 enterobacte
8	342.5	14.5	297	2 Q9Z349	Q9Z349 campylobact
9	246.5	10.4	355	2 Q9Z5241	Q9Z5241 helicobacte
10	240.5	10.2	355	2 Q9ZLX5	Q9ZLX5 helicobacte
11	123	5.2	278	2 Q9XB53	Q9XB53 erwinia car
12	117	5.0	1686	4 Q00443	Q00443 homo sapien
13	111.5	4.7	602	11 Q63485	Q63485 rattus norv
14	108.5	4.6	821	2 Q59241	Q59241 bacillus sp
15	107.5	4.5	824	2 Q9F216	Q9F216 bacillus sp
16	106.5	4.5	901	5 Q18749	Q18749 caenorhabdi
17	106	4.5	798	4 Q9UN32	Q9UN32 homo sapien
18	106	4.5	798	4 Q9UBK2	Q9UBK2 homo sapien
19	103	4.4	435	2 Q9X9C0	Q9X9C0 streptococc

Q9V9E6 drosophila
O84941 streptococc
P91736 hydra magni
Q45554 bacillus sp
Q9K48 bacillus ha
Q61194 mus musculu
Q61182 mus musculu
Q65553 bovine herp
Q9rdw6 lactobacilli
Q9rcz6 streptomyc
Q9p25 ureaplasma
Q9ujm5 homo sapien
Q9bty2 homo sapien
Q9scv5 arabidopsis
Q9c815 arabidopsis
Q9a6u7 caulobacter
Q59154 anaerocellu
O49137 arabidopsis
Q19319 caenorhabdi
Q91115 streptomyc
Q9z3v2 pseudomonas
Q9f192 arabidopsis
Q9h856 homo sapien
O54183 streptomyc
Q9y5c3 homo sapien
Q9y5f7 homo sapien

ALIGNMENTS

RESULT 1
Q9K0U7 PRELIMINARY; PRT; 382 AA.
AC Q9K0U7;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]

RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002403; AAF40901.1; -.
DR TIGR; NMB0464; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 34.5%; Score 815; DB 2; Length 382;

Best Local Similarity 44.7%; Pred. No. 8.7e-63;

Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSALNQDIMRLACVDTLVHGETPAVI-----KTKRSRLDDETIWQTI-KGPKQVYVQE 140

DB 33 LQCAALTNDVTRLACVDRIFAALPSSAGQEGESKAVLNLTETVRSSLDKGEAVIWEK 92

Qy	141	TTDDIFILMGNEKMLTKKDAKQLEYAAKQTFPLUSLFSFDLDRNN-TPLWSSRRPHNPMPVLP	199
Db	93	GGDAL-----PADSAGETADITYTPLSLMYDLDKNDLRLGLLVGREHNPMLP	139
Qy	200	IFVHGKFNRSPTPSH-EAKQFTPEFRAPELKFOYSVKVKAEDLWGTDSLWFCYTOQ	258
Db	140	LWYNNSPNYAPGSPTRGTTVOEKFQQKRAETKLQVSFKSIAEDLFKTRADLWFGYTOR	199
Qy	259	SHWOIFN-GRNSRPFRRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMGAHVHSGESAKLSR	317
Db	200	SDMQIYNQGRKSAPFRNTDKYPEIFLTQPVKADLPFGGRLRLMGVGFVHQSNGQSRPESR	259
Qy	318	SWNRAYLMAGWEKNLTVMPRIINGRIKFGSGSQPDNDPDLDYGYGDVRFYQLLENKS	377
Db	260	SWNRIFYAMAGMEWGKLTIVIPRVVWRAFDQ-SGDK-NDNPDADYMGYGQVKLYRLNDRQ	317
Qy	378	NISCTVYRNPSCGKALQLDYVYPLGKGISGYFOIFQGYGOSLIDYNHEATSFVGVLMLN	437
Db	318	NVYSVLRYNPKTGVAEEAYTFPIKGLKGVVGRFHYGESLIDYNHKKONGIGIGLMFN	377
Qy	438	DWMLG 442	
Db	378	DLDCI 382	
RESULT 2			
Q9JUT21	ID	PRELIMINARY;	PRT; 409 AA.
AC	Q9JUT21		
DT	01-OCT-2000 (TEMBLrel. 15, Created)		
DT	01-OCT-2000 (TEMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TEMBLrel. 17, Last annotation update)		
DE	PUTATIVE PHOPHOLIPASE.		
GN	NMA2021.		
OS	Neisseria meningitidis (serogroup A).		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=65699;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;		
RC	MEDLINE=20222556; PubMed=10761919;		
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,		
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,		
RA	Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,		
RA	Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,		
RA	Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,		
RA	Whitehead S., Spratt B.G., Barrall B.G.;		
RT	*Complete DNA sequence of a serogroup A strain of Neisseria		
RT	meningitidis Z2491.;		
RL	Nature 404:502-506(2000).		
DR	EMBL: AL162757; CAB85240.1; -.		
DR	InterPro: IPR003187; PLAI.		
DR	Pfam: PF02253; PLAI; 1.		
DR	Complete proteome.		
SW	SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;		

	Query Match	34.5%	Score 815;	DB 2;	Length 409;
	Best Local Similarity	44.7%;	Pred. No. 9.6e-63;		
	Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;
Qy	87	INCSALNODIMRLACACTYTLVHGCTPAVI-----KTKRSIRLDETIWQFI-KGKQVYVQE	140		
Db	60	LQCAALTDNVTFLACYDRIFAAQLPSSAGQGQSKAVLNTETVRSLLDGEAVIVVEK	119		
Qy	141	TTDPFIPLMGNEKGMILTKKDAKQLEYAAKQFTPLSLSFOLDLRRN-TPWLSSRPHNPMYVLP	199		
Db	120	GGDAL-----PADSAGETADIYTPLSIMYDLKNDLRLGLLVREHNPMYLMP	166		
Qy	200	IFMIGKPNRSPNTPSH-EAKQFTNEFRAPELKQVSVKVVAAEDLWCTDSDLNFGYTTQ	258		
Db	167	LWYNNSNPAGSPTRGTTVGEKEGQQRAETKLQVFSKIAEDLFTKRLADLWFGYTOR	226		

Qy	259	SHWOIFN-GKNSRFRVHDYQPEITFLTPQVYSDLPWDQKVRMTGMGVAHHSNGESA	KLRS	317
Db	227	SDWQIYNQGRKRSAPFRNTDYKPEITFLTPQVKADLPFGGRLRLMLGAGFVHQSGQSRPESR		286
Qy	318	SNWRAYLWAGMEKMLNLTVMPRINGRIKFEKSGSQDDNPDLIDYGYGDVRELYQLENKS		377
Db	287	SNWRIYAWAGMEWGLKLTIVPRVWVRAFDQ-SGDK-NONPDIADTMGVDYKLOVRLMDRQ		344
Qy	378	NISGTVRYNPSRSGKALQLDLYVYPPLGKISGYEQIFQGYQSLSLDYNHEATSFVGGLMLN		437
Db	345	NVYSVLRYNPKTGYGATEAAYTPPIKIKLKGVRGPHGYGESLIDYHNKQNGIGIGLMEFN		404
Qy	438	DWNGL	442	
Db	405	DLDGI	409	
RESULT 3				
Q9CL22				
ID	Q9CL22	PRELIMINARY;	PRT;	306 AA.
AC	Q9CL22;			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)			
DE	HYPOTHETICAL PROTEIN PM1426.			
GN	PM1426.			
OS	Pasteurella multocida.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;			
OC	Pasteurella.			
OX	NCBI_TaxID=747;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=PM70.			
EX	MEDLINE=21145866; PubMed=11248100;			
RA	May B.J., Zhang Q., Li L.B., Paustian M.L., Whittam T.S., Kapur V.;			
RT	"Complete genomic sequence of Pasteurella multocida PM70.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).			
DR	EMBL; AE006179; AA03510.1; -			
DR	InterPro; IPR003187; PLAL.			
DR	Pfam; PF02253; PLAL; 1.			
KW	Hypothetical protein; Complete proteome.			
SQ	SEQUENCE 306 AA; 35580 MW; EAF3DE8CLC22B26E CRC64;			
Query Match				
Best Local Similarity		15.7%	Score 372; DB 2; Length 306;	
Matches 83; Conservative		40; Mismatches 82; Indels 8; Gaps		

[illegible]

DE		PHOSPHOLIPASE A (EC 3.1.1.32).	
GN	PUDA OR CJIJ35I.		
OS	Campylobacter jejuni.		
OC	Bacteria; Proteobacteria;	epsilon subdivision; Campylobacter group;	
CC	Campylobacter.		
OX	NCBI_TaxID=197;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCTC 11168;		
RX	MEDLINE=20150912; PubMed=10688204;		
RA	Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,		
RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,		
RA	Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,		
RA	Quail M.A., Rajandream M.G., Rutherford K.M., Van Vleet A.H.M.,		
RA	Whitehead S., Barrell B.G.;		
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni		
RT	reveals hypervariable sequences.";		
NL	Nature 403:665-668(2000).		
DR	EMBL: AL139078; CAB73778.1; -.		
DR	InterPro; IPR003187; PLAL.		
DR	PFAM; PF02253; PLAI; 1.		
DR	PRINTS; PR01486; PHPLPASEA1.		
KW	Complete proteome.		
SQ	SEQUENCE 329 AA; 38880 MW; A32731F2B751AC44 CRC64;		
<div>Query Match 15.5%; Score 366; DB 2; Length 329; Best Local Similarity 33.2%; Pred. No. 9.2e+24; Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps</div>			
QY	130 IKGKPQQVYQTDPFLMGNEKGMTKKDAKOLEYAAKQTPLSLSFDLDNRNT--PLM 187 : : : : : : : :		
Db	43 LKNSSVLISQEONSSOATOTQSITIKKEEQ-----DFSRLALANYLGENSEFNPLG 96 : : : : : : : :		
QY	188 SSRPNHMYLVLPFMHGCKNPSPNTSYHEAKQFTPFEFAPELKFOVSVKVAADWLGT 247 : : : : : : : :		
Db	97 IS-SYKMNYELP-FAYSFNSGLVNNNKSEA-----KFOLS VKKRLLFENLLGL 141 : : : : : : : :		
QY	248 DSLDLFGYTQOSHWAQFNKGNSRPPFRVDHYQPFIETPTQPVY-SDLPWDGKVRMIGMAVGAV 306 : : : : : : : :		
Db	142 DEKYITYATTQSNWIY-EHS SPRETNYQPEFFIDLPLYLKDYEFNNLR---VGILH 196 : : : : : : : :		
QY	307 HSGESAK--LSRSWNRAYLMAGMEWKNLTYMPRIWGIRIFKGGSGSQPDNDPDILDYYGY 364 : ; : : : : : : :		
Db	197 ESNGKGDLENQRSWRNIYSTAILYNKFLFPVRLWTRI---PENKKDDDNPAILLHMGN 253 : ; : : : : : : :		
QY	365 GDVRFXYOLEKNKSINSTVRYPNR--SGKCALOLDIVYPL-GKGISGYFQIQGYCOSLI 421 : : : : : : : : : : :		
Db	254 FDVNLAY-LGDDYFILMLRNKLKFHNNGKAIOQLDGDFINFNGIIYWYLQYFNPGYGESLI 312 : : : : : : : : : :		
QY	422 DYNHEATSFCGVGLMLN 437 : : : : : : : : : :		
Db	313 DYKNHLQRSLTGFLIS 328 : : : : : : : : :		
RESULT 5			
ID	Q9SID7 PRELIMINARY; PRT; 292 AA.		
AC	Q9SID7;		
DT	01-MAY-2000 (TREMBlrel. 13, Created)		
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)		
DE	01-JUN-2001 (TREMBlrel. 17, Last annotation update)		
DE	PHOSPHOLIPASE A.		
GN	PIDA.		
OS	yersinia pseudotuberculosis.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
CC	yersinia.		
OX	NCBI_TaxID=633;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=YPIII PIB1;		
RA	Karlyshev A.V., Winzeler E.A., Williams K.J., Oyston P.C.,		
RA	Titball R.W., Wren B.W.;		

```

Qy 345 KEGSQDDNDPILDYGYGDRFLYQLENKSNISGTVRNPRSGKALQDYVYPLGK 404
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 -----GSTDDNDPDKYGYQKGYHL-GEAVLSAKGYNWNTGYGAEVGLSPYTK 253

Qy 405 GISGYFQIFQGYGSLIDYNHEATSFVGLMLND 438
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 254 HVRLYTVQYSGYGESLIDYNFQTRVGVGVMLND 287

RESULT 7
ID Q924N8 PRELIMINARY; PRT; 292 AA.
AC Q924N8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (EC 3.1.1.32).
GN PLDA.
OS Enterobacter agglomerans.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G., Brinkman E., van Boxel R., Bekkers A.C., Verheij H.M.,
RA Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120145; PubMed=9921577;
RA Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tommassen J.;
RT "Sequence comparison of outer membrane phospholipases A: implications
RT for structure and for the catalytic mechanism.";
RL Res. Microbiol. 149:703-710(1998).
DR EMBL; AF034414; AAD03498.1; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL.
DR PRINTS; PR01486; PHPLIPASEAL.
KW Signal; Hydrolase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 292 OUTER MEMBRANE PHOSPHOLIPASE A.
FT SEQUENCE 292 AA; 33719 MW; B75516D093B2BEEA CRC64;

Query Match 15.1%; Score 356.5; DB 2; Length 292;
Best Local Similarity 40.1%; Pred. No. 5.2e-23;
Matches 85; Conservative 31; Mismatches 87; Indels 9; Gaps 5;

Qy 229 ELKQVSVKVAEDLWCTDSLWFGYTOQSHWQIFNCKNSRPFVRVHDYQPEIFLTQPV- 287
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 86 EVKFQSLGFFPIWGAGDNLGASYTORSMWQASNDSSPFRETNYEPOIFLAWMTD 145

Qy 288 YSDLPWGDKVRMIGMGAHVHSGESAKLSRWNRAYLMAGMEWKNLTVMPIRWGRIFREG 347
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 146 YELAGW--TFREVEGYNHQSNRDPTRSRNDVYTRLMAQRGNLEIDLKFWYRIPESD 203

Qy 348 SGSPDDNDPILDYGYGDRFLYQL-ENKSNISGTVRNPRSGKALQDYVYPLGKI 406
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 S---KDDNPDINKYGYRLKVGALGESVFSIDG--RYNNTGYGGAEMGWSYPIKRV 258

Qy 407 SGYFQIFQGYGSLIDYNHEATSFVGLMLND 438
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 RFTYQVSGYGESMIDYNFQTRVGVGVMLND 290

RESULT 8
ID Q32349 PRELIMINARY; PRT; 297 AA.
AC Q32349;

```

```

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Campylobacter coli.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=195;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=UA585;
RC MEDLINE=97230284; PubMed=9119448;
RA Grant K.A., Belandia I., Dekker N., Richardson P.T., Park S.F.;
RT "Molecular characterization of plda, the structural gene for a
RT phospholipase A from Campylobacter coli, and its contribution to cell-
RT associated hemolysis.";
RL Infect. Immun. 65:1172-1180(1997).
DR EMBL; Y11031; CAA71915.1; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL.
DR PRINTS; PR01486; PHPLIPASEAL.
SQ SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;

Query Match 14.5%; Score 342.5; DB 2; Length 297;
Best Local Similarity 32.4%; Pred. No. 8.9e-22;
Matches 97; Conservative 37; Mismatches 108; Indels 57; Gaps 11;

Qy 178 DLDNRNTPWLSRRPHNMYVLPIFMHGKPNRSPNT-----P 213
      || || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16 DLKENNASLLSRKHETQ-----NTQKTPSTKEDFSRIALANYLGENSEFNPLGIS 66

Qy 214 SHEAKQETP-----NEFRAPELKFOVSVKVAEDLWCTDSLWFGYTOQSHWQIF 264
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 SYKMNYPFFAYSGSLGGENRKTENKFKLSIKRKFEDLLGLGKGYVGYTQTSWQ-- 124

Qy 265 NGKNSRPFVRVHDYQPEIFLTQPV-YSDLPWGDKVRMIGMGAHVHSGESAK--LSRSWNR 321
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 NYKHSSPFRETNYQPEFFVDIPLHEDYKFLNLR--VGILHESNGKDNLESRSNR 181

Qy 322 AYLMAEMWKNLTVMPIRWGRIFREGSGSQDDNDPILDYGYGDRFLYQLENKSNISG 381
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 IYASSVFLYQRFLEVPRIWYRI---PENSEDDNDNPEITHYMGNFEDIN-IGSLGNDYFNL 237

Qy 382 TVRYNP--RSKGKALQDYVYPL-CKGISGYFOIFQGYGSLIDYNHEATSFVGLMLN 437
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 238 MLRNLDPHDNKGAVQVDIGDIFDNGIYWIYQYENGIGGSLIDINKRLQLRLTAFLIS 296

RESULT 9
ID O25241 PRELIMINARY; PRT; 355 AA.
AC O25241;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).
GN HP0499.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RC MEDLINE=97394467; PubMed=9252185;
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.M.,

```

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*."
RL Nature 388:539-547(1997).
DR EMBL: AE000564; AAD07564.1; --.
DR TIGR: HP0499; --.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 2.7e-13;
Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KDKAKOLEYAAKQFTPLSLSFDLDRNNTPLWSSRPHPNMYVLPIFMHGKPNRSPNTPPSHE 216
DB 69 KKYLNMDYLGTYFLPFYHSF-----TFIQWYHPNINP----- 102
QY 217 AKOFTNEFRAPELKFOVSQVKKAAEDLWGTDSDLWFGYTQOSHWOIFNGKNSRPRVRHD 276
DB 103 ---YQRNEF-----KFOISFRVPVFRHILWTKTGLYLAITQDWFQIYNDSOPAPMRMN 154
QY 277 YOPEIFLTPQVSDLPDQGV---RMIGMAVHHSNG-ESAKLSRSWNRAYLMAGMEWKN 332
DB 155 FPELIYVYPI-NFKPFGKGKIGNFSEIWIGWQHISNGVGAOCYQPFNK-----EGNPENQ 209
QY 333 LTVMPRI-----WGRIFKEGSGSQP-----DDNPD 357
DB 210 FPGQPVIVDYNGQKQVWRGGRSVSAGQPVFRLWKEGGLKIMVAYWPYVPYDQSNEN 269
QY 358 ILDYGYGQVDFLY-----OLENKSNTSGTVRNPYNSRSGKALQLDYVPLGKIGISGF 410
DB 270 LIDIMYGNKIDYRGRHHEFLQYDIFQYWRD--RWHGAFRLGYTYRINPFVGIYA 327
QY 411 QIFQGYGQSLIDYNEATSFYGLMLN 437
DB 328 QWFNGYDGLYEYDFNSRIGVGIRLN 354

RESULT 10
Q9ZLX5 PRELIMINARY; PRT; 355 AA.
AC Q9ZLX5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE PHOSPHOLIPASE A1.
GN PLDA OR JHP0451.
OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).
OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
OC *Helicobacter*.
OX NCBI_TaxID=85963;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99120557; PubMed=9923682;
RX Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen *Helicobacter pylori*."
RL Nature 397:176-180(1999).
DR EMBL: AE001479; AAD06029.1; --.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KW Complete proteome.
SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 8.9e-13;
Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;
QY 157 KDKAKOLEYAAKQFTPLSLSFDLDRNNTPLWSSRPHPNMYVLPIFMHGKPNRSPNTPPSHE 216
DB 69 KKYLNMDYLGTYFLPFYHSF-----TFIQWYHPNINP----- 102
QY 217 AKOFTNEFRAPELKFOVSQVKKAAEDLWGTDSDLWFGYTQOSHWOIFNGKNSRPRVRHD 276
DB 103 ---YQRNEF-----KFOISFRVPVFRHILWTKTGLYLAITQDWFQIYNDSOPAPMRMN 154
QY 277 YOPEIFLTPQVSDLPDQGV---RMIGMAVHHSNG-ESAKLSRSWNR----- 321
DB 155 FPELIYVYPI-NFKPFGKGKIGNFSEIWIGWQHISNGVGAOCYQPFNKGNPENQFPQG 213
QY 322 -----AYLMAG-----MEWK--NLTVMPRINGRIFKEGSGQDD- 354
DB 214 PVTIKDYNGQKQVWRGGRSVSAGNALCFVLWKEGGLKIMVAYWPYV-----PYDQ 265
QY 355 -NPDIIDLYGYGQVDFLY-----OLENKSNTSGTVRNPYNSRSGKALQLDYVPLGKGI 406
DB 266 SNPOLIDYGYGNKIDYRGRHHEFLQYDIFQYWRD--RWHGAFRLGYTYRINPFV 323
QY 407 SGYFQIFQGYGQSLIDYNEATSFYGLMLN 437
DB 324 GIYAQWFNGYDGLYEYDFNSRIGVGIRLN 354
RESULT 11
Q9XB53 PRELIMINARY; PRT; 278 AA.
AC Q9XB53;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE KDU1.
GN *Erwinia carotovora*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCBI_TaxID=554;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-ATCC 39048, GS101;
RX MEDLINE=98065591; PubMed=9402024;
RA McGowan S.J., Sebaihia M., O'Leary S., Hardie K.R., Williams P.,
RA Stewart G.S., Bycroft B.W., Salmund G.P.;
RT "Analysis of the carbenepem gene cluster of *Erwinia carotovora*:
RT definition of the antibiotic biosynthetic genes and evidence for a
RT novel beta-lactam resistance mechanism."
RL Mol. Microbiol. 26:545-556(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 39048, GS101;
RX MEDLINE=98276484; PubMed=9614345;
RA McGowan S.J., Bycroft B.W., Salmund G.P.;
RT "Bacterial production of carbenepems and clavams: evolution of beta-
RT lactam antibiotic pathways."
RL Trends Microbiol. 6:203-208(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 39048, GS101;
RA McGowan S.J.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: U17224; AAD38237.1; --.
SQ SEQUENCE 278 AA; 31341 MW; 63769F4F3550E1B7 CRC64;

Query Match 5.2%; Score 123; DB 2; Length 278;
Best Local Similarity 21.6%; Pred. No. 0.011;
Matches 74; Conservative 52; Mismatches 128; Indels 88; Gaps 16;


```
Query Match      4.6%; Score 108.5; DB 2; Length 821;
Best Local Similarity 19.9%; Pred. No. 0.94;
Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;
```

Db	9	LISSILLIVLLSL-----PPTAAGCNTREDNFKHLLGNDVNRKSEAGALQLOQEV	63
Qy	65	ANPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMRLACY-DTLVHG	110
Db	64	QMTLVQGHGEKIQLRGSMSTHGLQWFFELNDNAYKALANDWESNMIRLAWYVGENGYAS	123
Qy	111	PAVTKRSIRLDETI-----WQT-----IKGKPOVV	137
Db	124	PELKSRVIRKIGIDLAIFENDMYVIVDMHVHAPDPRDPVYAGAEDFFRDIAALYPNNPHI	183
Qy	138	YQETDP-----IFLWGNKGLTKKDAKQLEAYAAKQFTPLSLSDLDNR-----NTPL	186
Db	184	YELANEFSSNNNGAGTPNNEEGHNAV-----EYADPIVEMLRDSGNADDNIIIVGSPN	238
Qy	187	WSSRP-----HNPMYVLPIF---MHGKPNRS---PNTPSHEAKOFTNEFRAPELKE	232
Db	239	WSQRPDLAANDPIDDHTMTVHFYTGSHAATESYPPETPNSEGRGNVMSNTRYA-----L	294
Qy	233	QVSVKVAADLWGT-----DSDLWFGYTOQ-----SHWQIFGNKS-----R	270
Db	295	ENGVAVFATE--WGTQOANGGGPYFDEADVWIEFLNENNISWANWSLTN-KNEVSGAFT	351
Qy	271	PFVRH-----NGDSPVEDVVIENEGALKLSGLDASNDVSEGNYNANRLSADGW	462
Db	352	PFEUGKSNATSLDGPDPQVWVPELSUGEVKVRIRKGVNTEP---IDRTKYTKVLWDFN	408
Qy	294	DGKVRMTGMAVHSHNGES-----AKLSRSMNRAYLMAGMEWKNLTVMPIRW	340
Db	409	DGTQKGFV-----NGDSPVEDVVIENEGALKLSGLDASNDVSEGNYNANRLSADGW	462
Qy	341	GR 342	
Db	463	GK 464	

RESULT 15

Q9F216 PRELIMINARY; PRT; 824 AA.

AC Q9F216; STRAIN=KSM-S237;

DT 01-MAR-2001 (TEMBLrel. 16, Created)

DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)

DE CELLULOSE.

OS Bacillus sp.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI_TaxID=1409;

[.]

RP SEQUENCE FROM N.A.

RC STRAIN=KSM-S237;

RC MEDLINE=21036886; PubMed=11193393;

RA Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S., Kobayashi T., Ito S.;

RA "deduced amino acid sequence and possible catalytic residues of a thermotable, alkaline cellulase from an alkaliphilic Bacillus strain";

RL Biosci. Biotechnol. Biochem. 64:2281-2289(2000).

DR EMBL: AB018420; BAB19360.1; -

DR InterPro: IPR001547; glyco_hydro_F5.

DR Pfam: PF00150; cellulase; 1.

DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.

SQ SEQUENCE 824 AA; 91564 MW; 65FA940FE1D729B9 CRC64;

Qy	30	PVAFVDEVRSEND-----LGQDNELPIDVOSATQSASTDTANPL-DEHEPELY-----TT	78
Db	25	PAALAAGCNTREDNFKHLLGNDVNRKSEAGALQLOQEVGQMTLVQGHGEKIQLRGSMSTH	84
Qy	79	ALENKTMLINCSAL-----NODIMRLACYDTLVHGE-----TPAVIKTK-----RSI	120

Query Match 4.5%; Score 107.5; DB 2; Length 824;

Best Local Similarity 19.7%; Pred. No. 1.2;

Matches 91; Conservative 56; Mismatches 147; Indels 167; Gaps 27;

```
Db 85 GLQWPEILNDNAYKALSNDSDNNIRLAMY-----VGENGYATNPelikQVRVIDGIELAI 140
Qy 121 RLDETI---WQT-----IKGKQVYQETTD-----IF 146
Db 141 ENDMYVIVDWHVHAPGDPDPVYAGAKDFFREIAALYPNNPHIYYELANEPSSNNNGGAG 200
Qy 147 LMGNEKGLTKKDAKQLEYAAKQFTPLSLSPDLDRN-----NTPLMSSRP-----H 192
Db 201 IPNNEEGKAVK-----EVADPIVEMLRKSGNADNIIIVGSPNWSQRPDLAADNPIDH 255
Qy 193 NPMYVLPITFMHGKPNRSPNTPSHEAKQFTNPEFRAPELK-----FQSVKVKAAEDLWGT 247
Db 256 HTMYTVHFYTGSHAATESYPSE-----TPNSERGNVMSNTRYALENGVAVFATE--WGT 308
Qy 248 -----DSDLWFGYTOQ-----SHWOIFNGKNS-----RPFV----- 274
Db 309 SQASGCGGPFDEADVWIEFLNENNISWANNSLTN-KNEVSGAFTPFELGKSNATNLDPG 367
Qy 275 --HDYQP-EIFLT-----QPV---YSDLPW---DGKVRMIGMGAVHHSNGE 311
Db 368 PDHVWAPEELSLSGEYVRARIKGVNYEPIDRTKYTKVLWDFNDGFKQ--GFGVNSDSPNK 425
Qy 312 SAKLSRSNRAYLMAGME-----WKNLTVMPRIWGR 342
Db 426 ELIAVDNENNTLKVSGLDVSNVDGDNFWANARLSANGWGK 466
```

Search completed: November 30, 2001, 14:26:30
Job time: 567 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:27:05 ; Search time 28.02 Seconds
(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 2363

Sequence: 1 MKVSLSTLTLSILPCFAILA.....YNHEATSGVGLMLNDWMGL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 10059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	373	15.8	1 P11_KLEPN	P37446 klebsiella
2	358	15.2	1 P11_SALTY	P37442 salmonella
3	357	15.1	1 P11_ECOLI	P00631 escherichia
4	321.5	13.6	1 P11_PROVU	P37447 proteus vul
5	108.5	4.6	1 GUN_BACS1	P05564 bacillus sp
6	98.5	4.2	1 ANDM_YEAST	P15274 saccharomyc
7	95	4.0	1 YJH0_YEAST	P40361 saccharomyc
8	94.5	4.0	1 GUN_PAEPO	P23548 paenibacilli
9	94.5	4.0	1 BCK1_YEAST	P01389 saccharomyc
10	93.5	4.0	1 SRK6_BRAOL	Q09092 brassica ol
11	92.5	3.9	1 VGI4_BPML5	Q05220 mycobacteri
12	92.5	3.9	1 BAC2_MOUSE	P97303 mus musculu
13	92	3.9	1 YL70_ARCFU	O28112 archaeoglob
14	92	3.9	1 R114_HUMAN	P48552 homo sapien
15	91	3.9	1 CATA_DICDI	O77229 dictyostell
16	91	3.9	1 YP72_CAEEL	Q09221 caenorhabdi
17	90.5	3.8	1 INVO_MOUSE	P48997 mus musculu
18	89.5	3.8	1 PBPA_STRPN	Q04707 streptococ
19	89.5	3.8	1 Y075_MYCGE	P47321 mycoplasma
20	89.5	3.8	1 MET_MOUSE	P16056 mus musculu
21	89	3.8	1 NPKE_BACAM	P06832 bacillus am
22	88.5	3.7	1 DCUP_BACSU	P32395 bacillus su
23	88.5	3.7	1 YJTH_HAEIN	P44544 haemophilus
24	88.5	3.7	1 CD44_RAT	P28051 rattus norv
25	88.5	3.7	1 AMY_ALTHA	P29957 alteromonas
26	88.5	3.7	1 GUN3_BACS4	P19570 bacillus sp
27	88.5	3.7	1 SABC_STRSL	Q55242 streptococ
28	88	3.7	1 MURE_CHLTR	O84271 chlamydia t
29	88	3.7	1 DTXH_CORBE	P00589 corynebaph
30	87.5	3.7	1 SEIL_HUMAN	Q9ubv2 homo sapien
31	87	3.7	1 LPFC_SALTY	P43662 salmonella
32	87	3.7	1 DP3A_BACSU	O34623 bacillus su
33	87	3.7	1 A2MG_HUMAN	P01023 homo sapien

ALIGNMENTS

RESULT 1

ID	P11_KLEPN	STANDARD;	PRT;	286 AA.
AC	P37446;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PLDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
OX	NCBI_TaxID=573;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94131966; PubMed=8300539;			
RA	Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.;			
RT	"Molecular characterization of enterobacterial plidA genes encoding			
RT	outer membrane phospholipase A.;"			
RL	J. Bacteriol. 176:861-870(1994).			
CC	- - FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	- - CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- - CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- - COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	- - SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X76901; CAA54223.1; -
PIR; B36971; B36971.
PIR; S40129; S40129.
DR InterPro: IPR003187; PLAL.
DR Pfam: PF02253; PLAL; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20
FT CHAIN 21 286 PHOSPHOLIPASE A1.
FT ACT_SITE 161 161 BY SIMILARITY.
SQ SEQUENCE 286 AA; 32544 MW; 3E39F863085108A3 CRC64;

Query Match 15.8%; Score 373; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 7.4e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 30-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=85003590; PubMed=6383820;
RA de Geus P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
RT "The pro- and mature forms of the E. coli K-12 outer membrane
RT phospholipase A are identical.";
RL EMBO J. 3:1799-1802(1984).
RN [7]
RP SEQUENCE OF 174-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87115164; PubMed=3027506;
RA Iriano N., Nakayama K., Nakayama H.;
RT "The recQ gene of Escherichia coli K12: primary structure and
RT evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).
RN [8]
RP MUTAGENESIS OF SER-172.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A";
RL J. Bacteriol. 176:861-870(1994).
RN [9]
RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
RX MEDLINE=91249806; PubMed=2040286;
RA Horrevorts A.J.G., Verheij H.M., de Haas G.H.;
RT "Inactivation of Escherichia coli outer-membrane phospholipase A by
RT the affinity label hexadecanesulfonyl fluoride. Evidence for an
RT active-site serine.";
RL Eur. J. Biochem. 198:247-253(1991).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
CC DORMANT IN NORMAL GROWING CELLS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
CC LYSIS OR TEMPERATURE SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02143; CAA26081.1; -
DR EMBL; M87049; AAA67617.1; -
DR EMBL; AE000458; AAC76824.1; -
DR EMBL; AF005613; AAG59017.1; -
DR EMBL; AF002567; BAB38174.1; -
DR EMBL; M30198; AAA24516.1; -
DR PIR; A00771; PSECA.

DR PIR; A22133; PSECAL.
DR PIR; S30711; S30711.
DR EcoGene; EG10738; plda.
DR InterPro; IPR003187; PLAI.
DR Pfam; PF02253; PLAI; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
FT Complete proteome.
FT SIGNAL 1 20 PHOSPHOLIPASE A1.
FT CHAIN 21 289
FT ACT_SITE 164 164
FT MUTAGEN 172 172
FT CONFLICT 14 15 S->F: INACTIVE PROTEIN.
FT CONFLICT 30 33 LP -> FA (IN REF. 2).
FT SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CRC64;
Query Match 15.1%; Score 357; DB 1; Length 289;
Best Local Similarity 39.3%; Pred. No. 1.5e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;
QY 226 RAPELKFOVSVKVAEEDLWGTDSLWFGYTOQSHWQIENGKNSRPRVHDYQPEIFLTQ 285
DB 83 RDEVFQSLSLAPLWRGILGPNVSLGASTQKSWQLSNSESSPFRETNYEPQLFLGF 142
QY 286 PV-YSLDPWDGKYRMICMGAVHHSNGESAKLSRSNRAVLMAGMKWNLVTMPRIWGRIF 344
DB 143 ATDYREFAGW--TLRDVEMGVNHDNCGSDPTSRNRLYTRLMAENGWNLVEVKPWVW- 199
QY 345 KEGSGQPDNDPILDYGYGVRFYOLENKSNIQSVYRNPNSGKQALQLDVYVPLGK 404
DB 200 -----GNTDDNDPDITKMGYYQLKIGYHL-GDAVLAKAQYNNWTGYGAELGSLSPITK 253
QY 405 GISGYQIFQGYGOSLDYNHEATSFVGLMLND 438
DB 254 HVRLYQVYSGYGESLIDYNFNQTRVGVGVMLND 287
RESULT 4
PAL_PROVU STANDARD; PRT; 289 AA.
AC P37447;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RP [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A";
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02143; CAA26081.1; -
DR EMBL; M87049; AAA67617.1; -
DR EMBL; AE000458; AAC76824.1; -
DR EMBL; AF005613; AAG59017.1; -
DR EMBL; AF002567; BAB38174.1; -
DR EMBL; M30198; AAA24516.1; -
DR PIR; A00771; PSECA.

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X76902; CAA54224.1; -;
DR PIR; C36971; C36971.
DR PIR; S40130; S40130.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 20 BY SIMILARITY.
FT CHAIN 21 289 PHOSPHOLIPASE A1.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 289 AA; 32944 MW; D75516CFPB406997 CRC64;

Query Match 13.6%; Score 321.5; DB 1; Length 289;
Best Local Similarity 33.3%; Pred. No. 1e-18;
Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

Qy 179 LDRNNTPLWSSRPHPMVLPIFMHGKPNRSPTPSHEAKOFTPNEFRAPELKFQVSVKV 238
Db 43 LQEHDP-FTLYPESNXYLLTY-----TSLNKAIESYNWSDNA-NKDEVKFSLSLAF 95
Qy 239 KAAEDLWGTSDLWFGYQQSHWQIFNGKNRPPRVHDYQPEIFLTQPV-YSDLPWDGKV 297
Db 96 PLWRGILGNSLGSASYQTSWQWLSNTGESAPFETNYEPQLFGFATDYSVGDW--TL 153
Qy 298 RMIGMGAVHHNGESAKLSRSNNRAYLMAGMEWKNLTVMPRTWGRIFKEGSGQDDNPD 357
Db 154 RDAEFGYNHQNSRDPTRSRNRLYSRLMAQNGNLVVEKPVYI-----GDTSDNKN 207
Qy 358 ILDYYGYGDVRFLEYOLENKSNIISGTVRNPNSGKALQLDYVPLGKISGYFQIFQYGV 417
Db 208 ITKYMGGYQLKIGYQL-GEAVLSAKQYNWNTGYGGAELGVSGYPTTKHVRFTQVYSYG 266
Qy 418 QSLIDYNHEATSFYGLMLND 438
Db 267 ESLIDYDFNQTRVNGVMLND 287

RESULT 5
GUN_BACS1 STANDARD; PRT; 800 AA.
AC P06564;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE (ALKALINE CELLULASE).
OS Bacillus sp. (strain 1139).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1411;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87085443; PubMed=30989509;
RA Fukumori F., Kudo T., Narahashi Y., Horikoshi K.;
RT "Molecular cloning and nucleotide sequence of the alkaline cellulase
gene from the alkalophilic Bacillus sp. strain 1139.";
RL J. Gen. Microbiol. 132:2329-2335(1986).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
LINKAGES IN CELLULOSE.
CC -1- MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES
ONE ALKALINE CELLULASE (PH OPTIMUM 9.0). THIS STRAIN IS NOT A TRUE
CELLULOLYTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO
HYDROLYSE NATIVE CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M15743; AAA22305.1; -;
DR EMBL; D00066; BAA00045.1; -;
DR PIR; A29003; A29003.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 800 ENDOGLUCANASE.
FT ACT_SITE 190 190 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 305 305 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 800 AA; 88602 MW; 7CCA4D7B6DAD55CF CRC64;

Query Match 4.6%; Score 108.5; DB 1; Length 800;
Best Local Similarity 19.9%; Pred. No. 0.6;
Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

Qy 8 LTLSELPCFAILAIQAOAVPNPVAFVDEVREND---LGQDNELPIDVQSATOSASTDT 64
Db 10 LISSILILVLLSL-----FPTALAAEGNTREDNFKLLGNDNVKRPSEAGALQLEVDG 64
Qy 65 ANPL-DEHEPELY-----TTALENKTMLNCSAL-----NODIMRLACY-DTLVHGET 110
Db 65 QMTLVQDHGEKIQLRGSMTHGLQMPPEILNDNAYKALANDWESNNIRLAMYVGNGYASN 124
Qy 111 PAVIKTKRSIRLDETI-----WQT-----IKGKPOVV 137
Db 125 PELKSRVVIKIDIAIENDMTYIVDVHWHAPCDPRDPVYAGAEDEFRIALYNNPHII 184
Qy 138 YQETTPD-----IFLMGNEKGMITKKDAKQLEYAAKQFTPLSLSFOLDNRN---NTP 186
Db 185 YELANEPSSNNNGGAGIPNNEEGNAVK-----EVADPIVEMLRDSDGNADDDNIIIVGSPN 239
Qy 187 WSSRP-----HNPMYVLPF--MHGKPNRS--NTPSHEAKQFTPEFRAPELKF 232
Db 240 WSQRPDLAADNPIDDHHTMYTVHFYTGSHAASTESYPETPNSEGRGNVMSNTRYA-----L 295
Qy 233 QVSVKVAEADLWGT-----DSLWFGYTOQ-----SHWQIFNGKNS-----R 270
Db 296 ENGVAVFATE--WGTSQANGDGGPYFDEADVWIEFLNENNISWANWSLTN-KNEVSGAFT 352
Qy 271 PFRVH-----DYOEFLFQPVYSDLPW--- 293
Db 353 PFELGKSNATSLDPGPDQVWVPEELSLSGEVVRARIKGVNVEP---IDRTKYTKVLWDFN 409
Qy 294 DGKVRMIGMGAVHHSNGES-----AKLSRSNNRAYLMAGMEWKNLTVMPRIW 340
Db 410 DGTQKGFV-----NGDSPVEDVVIENAGALKUSLDASNDVSEGNVYANARLSADGW 463
Qy 341 GR 342
Db 464 GK 465

RESULT 6
ADNM_YEAST STANDARD; PRT; 810 AA.
AC P15274;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE AMP DEAMINASE (EC 3.5.4.6) (MYOADENYLATE DEAMINASE).
GN AMD1 OR AMD OR YML035C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;

Query Match
Best Local Similarity 4.0%; Score 93.5; DB 1; Length 849;
Matches 78; Conservative 21.9%; Pred. No. 10; Mismatches 50; Indels 121; Gaps 24;

QY 105 LVHGETPAVITKRSIRDETQITQIKQPVVYQETDPIFLMG-----NEK---GMLTK 157
DB 24 LIHPALSIYINT-----LSSTESLTISNNKTLV---SPGISFEVGFRTNSRWYLCGMVK 75
QY 158 K-DAKOLEVAAKQFTPLSLSPD-----LDNRNTPWSSRPHNPMYVLPFIMHGK 205
DB 76 KVSDDTYVWVNRDPLNSAIGTLKISGNNVLLDHSNKPVMWTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAKQFT-----PNEFRAPELKFQSVKVKAAEDLW 245
DB 126 -ERSPVVAELLANGFVNRDSSNNDASEYLWQSFDPDTDLLPENMKLGYNLT----- 177
QY 246 GTDSDLWFGYTOOSHWQIFNGKNSRFRVHDYQ-----PEIFLQO---PVYSDLPWDGK 296
DB 178 GLNREFL-----TSWRSDDPSSGNF---SYKLETQSLPEFLSRENFPMHRSGPWNG- 226
QY 297 VRMICMGAVH-----HSNGESAKLSRSWNRAY-----LMAGMEWKNLTVMV--RI 339
DB 227 IRFSGIPEDQKLSYVWYFNFIENNEVATFTNTNSFVRLTLISEGFQRLTWYPSIRI 286
QY 340 WGRIFKEGSGQPDNDPILDY-----YCYGDVRELYOLENKSNSGTVR--YNPRS 389
DB 287 WNREW-----SSPVD-PQCDTYMGCPAYCDV-----NTSPVCNCIQENPN 329

RESULT 11
VG14_BPML5
ID VG14_BPML5 STANDARD; PRT; 486 AA.
AC Q05220;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE GENE 14 PROTEIN (GP14).
GN 14.
OS Mycobacteriophage L5.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
OX NCBI_TaxID=31757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211282; PubMed=8459766;
RA Hatfull G.F., Sarkis G.J.;
RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
a phage system for mycobacterial genetics.";
RL Mol. Microbiol. 7:395-405(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z18946; CAA79390.1; .
DR PIR; S30959; S30959.
SQ SEQUENCE 486 AA; 53757 MW; 4BF3734AC6B11BF5 CRC64;

Query Match
Best Local Similarity 3.9%; Score 92.5; DB 1; Length 486;
Matches 99; Conservative 21.0%; Pred. No. 5.8; Mismatches 173; Indels 131; Gaps 24;

QY 21 IQQAQA-VPNVAFVDEVRSEN-----DLGQDNELPIDVQSATQSASTDTANPLDHEPE 74
DB 56 MQOLLAHVGYPRLYVDSVAERQAVEGFRGLGDADEAEELWQWQ-----ANNLDIEAPL 109
QY 75 LYTTALENKTMLINCSA-----LNQDIMRLUACVDTLVHGE-TPAVIKTKRSIRLD 123
DB 110 GYTDAYVHGSRSFITISKPPDQLDLGWDQNVPIRVE-PPTRMHAEIDPRINRVSKAIR-- 166
QY 124 ETIMWTIKGQVQVYQETDPIFLMGNEKGMTKKDAKQLEVAQAQFTPL-SLSFDLDRN 182
DB 167 -----VAYDKE-----GNE-----TOAATLYTPMETIGFRADG 195
QY 183 NTPWSSRPHNPMYVLPFIMHGKPNRSPNTPSHEAKQFTPNEFRAPELKFQSVKVKAAE 242
DB 156 EWAENFNYPHG-LGVVPV--PLPNRTRLSLDYGTSETPELRSMTDAAARILMLQATA 252
QY 243 DLWGTDSDLNF-----GYTOQSHWQIFNGKNSRFRVHDYQPEI-----F 282
DB 253 ELMGVPPQRLIFGIKPEEIGVDSETGQTLFDAYLARILAFEDAEGKIQOFSAAELANFTNA 312
QY 283 LTO-----PVYSDLPWDGKVRMIGMGAHVHNGES-----AKLSRSWNRAYLMAGMEWKN 332
DB 313 LDQIAKQVAAITGLP-----PQYLSTAADNPASAEAIRAAESRLKKVVERKNLMFGCAHEE 368
QY 333 LTVMPRIWGRIFKEGSGQPDNDPILDYGYGDVRELYOLENKSNSGTVRYNPRSGK 392
DB 369 AM---RIAYRIMKGG-----DVPPDML-----RME-----TVWRDPSTPTY 401
QY 393 ALOLD---YVYPLKGI--SGYFQIFQGYG-----QSLIDYNHEATSFVGVLG 435
DB 402 AAKADAATKLYGNGGQVIPRERARIDMGYSVKEREEMRRWDEEAAAGLGL 453

RESULT 12
BAC2_MOUSE
ID BAC2_MOUSE STANDARD; PRT; 716 AA.
AC P97303;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2).
GN BACH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=97042438; PubMed=8887638;
RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
RA Yamamoto M., Igarashi K.;
RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
RT transcription factors that interact with Mafk and regulate
RL transcription through the NF-E2 site.";
RL Mol. Cell. Biol. 16:6083-6095(1996).
CC -|- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
CC MAFK.
CC -|- SUBUNIT: HETERO-DIMER OF BACH2 AND MAFK.
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -|- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
CC NEURONAL CELLS.
CC -|- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
CC -|- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC


```
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: D86604; BAA13138.1; -
DR HSSP: P05412; IFOS.
DR MGD: MGI:894679; Bach2.
DR InterPro: IPR000210; BTB_POZ.
DR InterPro: IPR001871; BZIP.
DR Pfam: PF00651; BTB; 1.
DR Pfam: PF00170; BZIP; 1.
DR SMART: SM00338; BRLZ; 1.
DR SMART: SM00225; BTB; 1.
DR PROSITE: PS50097; BTB; 1.
DR PROSITE: PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; Repressor; DNA-binding;
KW Nuclear protein.
FT DOMAIN 37 103 BTB.
FT DOMAIN 162 168 POLY-GLU.
FT DNA_BIND 527 542 BASIC MOTIF.
FT DOMAIN 550 572 LEUCINE-ZIPPER.
SQ SEQUENCE 716 AA; 78935 MW; 9132B3731AE24333 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 716;
Best Local Similarity 22.8%; Pred. No. 10; Mismatches 127; Indels 149; Gaps 21;
Matches 93; Conservative 39;

QY 38 RSENDLG-----QNELPIDVQSATSQASTD-----TANPLDEH-----EPE 74
Db 151 RQEDHGNSAGEEEEEETMDSETARMACATDMLPDPISEFATAIPVAKEEALLPESE 210
QY 75 LYTALENKTMINSALNQ-----DIMRLAC-----YDTLVHG----- 108
Db 211 VPTDKENSEK-----GALTQYPRYKYLQACTKNYSAPSHSGTGFASFSDSPGNSLK 266
QY 109 -----ETPAVKTKRSRL-----DETWTQIKGPOVVYQETTDPI----- 145
Db 267 PGLPMGQIKSEPPSEETEEETLCLSGDET-----DIKDRPGDVEMDRKQPSPARTPTSTR 323
QY 146 -----FLMNEKGMLTKDKAKOLEYAAKQFTLSLSF----- 177
Db 324 GAACLDRSVSSPCLSLFGITGV-----ESTGLPSTSQ-----PLVRSACPPNKGISQ 377
QY 178 -DLDRNPTPLWSS--RPHNPMYVLPFIHMGKPNRSP-----NTPSHEAKQFTPNF 225
Db 378 GDLKDYTPLAGNYGQPHVQGVQDVSNFAMGSLRPGPGLTCLFESSQSCGARELATEH 437
QY 226 RAPELK-----FQSVKVKAAEDLNGTSDSLWFGYTTQSHWQIFNGKNSRPRVHDYQPE 280
Db 438 QEPGLMGDMYNOVQPIK-CEQSYGTNSSDESG-----SFSEADSESCPVQDRGOE 488
QY 281 IFLTQPV--YSDLPWDGKVRMIGMGAVHNSGESAKL-----SRSWNR 321
Db 489 VKLPFPVQDITDLPNDQFMIMKM---HKLTSQLEFTHDIRRSKNR 533

RESULT 13
YL70_ARCFU STANDARD; PRT; 324 AA.
AC O28112;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE HYPOTHETICAL PROTEIN AFE2170.
GN AFE2170.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
```

```
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Cocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: CONTAINS 2 KELCH REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: AE000955; AAB89093.1; -.
DR TIGR: AF2170; -.
DR InterPro: IPR001798; Kelch.
DR Pfam: PF01344; Kelch; 2.
KW Hypothetical protein; Repeat; Complete proteome.
FT REPEAT 229 276 KELCH 1.
FT REPEAT 277 323 KELCH 2.
SQ SEQUENCE 324 AA; 36025 MW; C445388CFEB96E45 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 324;
Best Local Similarity 20.1%; Pred. No. 3.7; Mismatches 72; Indels 98; Gaps 18;
Matches 54; Conservative 44;

QY 224 EFRAPELKQVSVKVKAAED-----LWGTSDSLWFGYTTQSHWQIF-NGKNSRPRVH 275
Db 63 EFHPPN-RLIEI---LLSSEDFSQREAANVDGEILLFGT-----VFENGKYSPTDQIL 112
QY 276 DYQPEIFLTQPVYSDLP-----W-DGKVRMI-----GMGAVHNSGESAKLSRS 318
Db 113 SFNPKLERLRVLNASLPHPTSDVAAVWGDVSRVYIFLNNSECEVAFYPSNESFAKLQVS 172
QY 319 -----W--NRAYL-----MAGMEKMLTVMPRIWGRIFKEGSGSQ 351
Db 173 CPTEHPGGCVHVVWYGGKAYFCGEGVASFDPMGGFKW--IAFTDRVWVRAATVADG-- 228
QY 352 PDNPDILDYGYGVDFRFLYQLENKSNISGT-----VRYNPRSGKALQLDYVYPLGKG-- 405
Db 229 -----YIFAIGSGSGIAETKDEIRFPNPKTGE-LCEMRKTLKLPVARGQA 270
QY 406 --ISG-YFOIF-----QGYQSLIDYNH 425
Db 271 VAVGGEYIYIFGGYTKDGYANEILRYDY 298

RESULT 14
RI14_HUMAN STANDARD; PRT; 1158 AA.
AC P48552;
DT 01-FEB-1996 (Rel. 33, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NUCLEAR FACTOR RI140 (NUCLEAR RECEPTOR INTERACTING PROTEIN 1).
GN NR1P1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

RC TISSUE-Breast;
RX MEDLINE=95369246; PubMed=7641693;
RA Cavalline V., Dauvois S., L'Horsset F., Lopez G., Hoare S.,
RA Kushner P.J., Parker M.G.;
RT "Nuclear factor RIPI40 modulates transcriptional activation by the
RT estrogen receptor.";
RL EMOB J. 14:3741-3751(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii Y., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Pooley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Leirach H., Reinhardt R., Vaspou M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- FUNCTION: MODULATES TRANSCRIPTIONAL ACTIVATION BY THE ESTROGEN
CC RECEPTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X84373; CAA59108.1; -;
DR EMBL; AL163207; CAB90396.1; -;
DR MIM; 602490; -;
KW Transcription regulation; Nuclear protein.
FT CONFLICT 124 124 P -> R (IN REF. 1).
FT CONFLICT 448 448 R -> G (IN REF. 1).
FT CONFLICT 721 726 NKGSE -> TKGKVK (IN REF. 1).
FT CONFLICT 1080 1080 T -> A (IN REF. 1).
SQ SEQUENCE 1158 AA; 126941 MW; 81FC424968E9A5F6 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 1158;
Best Local Similarity 20.0%; Pred. No. 21;
Matches 93; Conservative 62; Mismatches 138; Indels 172; Gaps 23;

Qy 41 NDLGODNLPIDVQSATQSASDTANPLDEHEPELYTTA-----LENKMTLIN----- 88
Db 524 NDVSKFNTQNTARTSVIESPSNTRTPVS--TPPLTSSKAGSPINLSQHSVLVKNWSP 581

Qy 89 --CSALNODIMRLACYDTL-----VHGETPAVTK-----TKRSIRLDETIWQTKGK 133
Db 582 YVCSTQSEKLTNTASNSMDLTKSDPPGKPAQNEGAQNSATFSASKLQNLQAQCGMQS 641

Qy 134 PQVYQETITDIFLMGNEKGMLTAKDAQLEYAAKQFTPLSLDFLDNRNTP----- 186
Db 642 SMSVEQRPSKQLTGN-----TDK-----PIGM-----IDRLNPLLSNKTNA 680

Qy 187 -----WSSRPHNP-----MYVLPTEHMGKPNRSPNTPSHKAKQFTP--- 222
Db 681 VEENKAFSSQTGPPEGLSGSEIENLLRRVLQLLL--GNPNKGS-----EKKEKTPURD 735

Qy 223 -----NEFRAPELKFQSVKVAEADLWGTDSLDLFGYTOQSHWQIFNGKNSRPF----- 272
Db 736 ESTQHSERALSQQLMKVYKSEPCDDIPIPTNV-----HLSHDAKSAAPFLGMA 785

Qy 273 -RVHDYQP-----EFLTPQVYSDLPWDGKVRMTGMGAVHNSGESAKLSRWNRYL-- 324

Db 786 PAVORSAPALPVSEDFKSEPV---SPQDFS-----FSKNGLLSLRLRQNDQSYLAD 833
Qy 325 -----MAGMEWKNLTVMPTWRIWGRIFKEGSGSQDDNP-----DILDYTG----- 363
Db 834 DSDSRHNRNEWALLESKNLCWVPK-----KRKLYTEPLENPFKMKNNIVDAANHSAP 887
Qy 364 ---YGDVRFYQLENKSNISCTVRYNPRSGKALQLDQYVYPLGKG 405
Db 888 EVLYGSLNLQNEELKFSRN-----DLEFKYPAGHG 916

RESULT 15
CATA_DICDI
ID CATA_DICDI STANDARD; PRT; 496 AA.
AC 077229;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA OR CAT.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Foote C., Alexander H., Alexander S.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF090443; AAC36743.1; -;
DR HSSP; P00432; 7CAT.
DR DictyDb; DD0722?; cata.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 54 54 BY SIMILARITY.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 338 338 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SITE 494 496 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 496 AA; 55683 MW; 683318B8FBAFD2E6 CRC64;

Query Match 3.9%; Score 91; DB 1; Length 496;
Best Local Similarity 23.5%; Pred. No. 7.9;
Matches 56; Conservative 32; Mismatches 94; Indels 56; Gaps 12;

Qy 61 STDANPLDEHEPELYTALENKTMNLNCSALNODIMRLACYDTLVHGETPAVTKRSI 120
Db 7 TTSGSGSPIDNN-----LNSMTAGVNGPLIQDTLI---DKLAHFDRE----- 46

Qy 121 RLDETIWQTKGKQVYVQETTD-----PIF-----LMGNEKGM-L 156
Db 47 RIPERVVHA-KGAGAHGYFEVTSDDVPKWKAKFLNKVGKFTPIFRFTVSGEGKSSDS 105

Search completed: November 30, 2001, 14:27:06
Job time: 553 sec

This Page Blank (uspto)